DOCUMENT RESUME

ED 274 231 HE 019 656

AUTHOR	Gray, Susan; And Others
TITLE	Validation, Edits, and Application Processing Phase II and Error-Prone Model Report.
INSTITUTION	Applied Management Sciences, Inc., Silver Spring, Md.
SPONS AGENCY	Office of Student Financial Assistance (ED), Washington, DC.
PUB DATE	80
CONTRACT	300-79-0742
NOTE	188p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	Accountability; College Students; *Disclosure; Eligibility; Error Patterns; Evaluation Criteria; Federal Programs; *Financial Aid Applicants; Grants; Higher Education; Management Information Systems; *Models; Program Administration; *Quality Control; Research Methodology; Statistical Analysis; *Student Financial Aid
IDENTIFIERS	*Basic Educational Opportunity Grant Program

ABSTRACT

The impact of quality assurance procedures on the correct award of Basic Educational Opportunity Grants (BEOGs) for 1979-1980 was assessed, and a model for detecting error-prone applications early in processing was developed. The Bureau of Student Financial Aid introduced new comments into the edit system in 1979 and expanded the pre-established criteria (PEC) used to select students for validation. Study objectives were to assess: the impact of the processing system edits on applicant behavior; the adequacy of the PEC as effective indicators of misreporting; and the impact of validation effort on the correct award of BEOGs. In analyzing the processing system edits, the focus was the frequency and type of corrections to key application items, and the ability of applicants to quickly meet the demands of the processing system and receive an eligibility determination. To determine the adequacy of PEC as indicators of misreporting, a sample of individuals selected for validation was compared with a randomly-selected sample. In assessing the impact of validation edits, a sample of validation applicants was compared with a similar sample of non-validation applicants. An error-prone model related to the validation system was also developed and compared with the existing PEC. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED274231

Prepared for:

Office of Student Financial Assistance
Department of Education

G-170

VALIDATION, EDITS AND APPLICATION PROCESSING PHASE II AND ERROR-PRONE MODEL REPORT

July 3, 1980

In Response to:

Contract No. 300-79-0742

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

HE 019 656

APPLIED MANAGEMENT SCIENCES

BEST COPY AVAILABLE

G-170

QUALITY CONTROL ANALYSIS OF SELECTED ASPECTS OF PROGRAMS ADMINISTERED BY THE BUREAU OF STUDENT FINANCIAL ASSISTANCE

This report was prepared pursuant to Contract No. 300-79-0742. The names of the persons employed by the contractor with managerial or professional responsibility for such work are as follows:

Ms. JoAnn Kuchak, Project Director

Ms. Susan Gray, Senior Project Analyst
(Data Analysis and Report Preparation)

Mr. John Neely, Project Analyst
(Data Analysis and Report Preparation)

Dr. Pedro Saavedra, Technical Manager
(Data Analysis and Report Preparation)

Applied Management Sciences
962 Wayne Avenue, Suite 701
Silver Spring, Maryland 20910

Office of Student Financial Assistance
Department of Education

TABLE OF CONTENTS

EXECUTIVE SUMMARY

CHAPTER — PAGE

1 STUDY OVERVIEW

1.1: Purpose and Scope. 1.1
1.2: Key Study Objectives and Questions 1.1
1.3: Study Design 1.3
1.4: Study Population 1.7
1.5: Study Methodology. 1.9
1.6: Study Caveats. 1.10

2 IMPACT OF THE PROCESSING SYSTEM EDITS AND COMMENTS

2.1: Global Impact of the Edits 2.2

2.1.1: Description and Distribution of Comments . . 2.2
2.1.2: Corrections in Response to Comments. 2.5
2.1.3: Current Eligibility and Payment Status of Applicants Receiving Comments 2.11
2.1.4: Summary and Conclusions. 2.14

2.2: Rejection Edits and Comments 2.15

2.2.1: Characteristics of Applicants Most Frequently Rejected. 2.15
2.2.2: Adequacy of the Individual Rejection Edits and Comments 2.21
2.2.3: Summary and Conclusions. 2.35

3 ADEQUACY OF THE PRE-ESTABLISHED CRITERIA

3.1: Background on the Pre-established Criteria 3.1

3.1.2: Study Design 3.3
3.1.3: Measures of Effectiveness. 3.4

3.2: Overall Effectiveness of the PEC as Compared to Random Selection 3.6

3.3: Relative Effectiveness of PEC Subgroups. 3.9

3.3.1: Rate of Applicants Making Post-Selection Corrections. 3.9
3.3.2: Average Effective SEI Changes. 3.11

TABLE OF CONTENTS
(Continued)

CHAPTER			PAGE
	3.3.3:	Composite SEI Changes.	3.13
	3.3.4:	Percent of Applicants Correcting Post-Selection with no SEI Change	3.13
	3.3.5:	Change to Critical Fields.	3.16
	3.3.6:	Relationship of Corrections Behavior of PEC Validation Applicants to "Suspect" Fields	3.19
	3.3.7:	Rate of Applicants Appearing on the Recipient File and Current Eligibility Status	3.22
	3.4:	Summary of Findings.	3.27
4		VALIDATION AND THE INTERACTION OF VALIDATION AND THE EDITS	
	4.1:	Impact of Validation - Comparison of Pre and Post-Selection/Eligibility Change.	4.2
	4.2:	Impact of Validation - Percent of Applicants Correcting Post Selection with No SEI Chance.	4.5
	4.3:	Impact of Validation - Applicants on (not on) the Recipient File and Their Current Eligibility Status.	4.6
	4.4:	The Interaction of the Impact of Validation and the Processing System Edits	4.9
5		ERROR-PRONE MODELLING STUDY	
	5.1:	Purpose and Scope.	5.1
	5.2:	Key Study Objectives and Questions	5.1
	5.3:	Study Population and Samples	5.2
	5.4:	Research Methodology	5.5
	5.5:	Dependent Variable	5.5
	5.6:	Predictor Variables.	5.7
	5.7:	Strengths and Limitations.	5.7
6		THE ERROR-PRONE MODEL	
	6.1:	Results of the Validation and Nonvalidation Sample Comparisons	6.1

TABLE OF CONTENTS
(Continued)

CHAPTER		PAGE
	6.2: Description of Model Results.	6.2
	6.3: Applicants Who Failed to Re-enter the System. . . .	6.10
	6.4: Misreporting by Critical Fields	6.11
	6.5: Impact of Misreporting.	6.14
7	DESCRIPTION AND INTERPRETATION OF THE GROUPS	
8	RECOMMENDATIONS AND EFFECTIVENESS OF THE MODEL	
	8.1: A Validation Selection Strategy in the Absence of Other Changes.	8.1
	8.2: Estimation of Taxes, An Alternative Remedy.	8.2
	8.3: Validation of Independents.	8.3
	8.4: The PEC Versus the EPM.	8.4
	8.5: Further Research Needed	8.4
	8.6: General Recommendations and Alternatives Related to the EPM.	8.5
	APPENDIX A: COMPUTER COMMENTS	
	APPENDIX B: ASSUMPTIONS FOR COMPUTING ELIGIBILITY INDEX WHEN APPLICANT IS REJECTED	
	APPENDIX C: PRE-ESTABLISHED CRITERIA	
	APPENDIX D: GENERAL METHODOLOGY FOR ERROR-PRONE MODEL	
	APPENDIX E: STEP-BY-STEP DEVELOPMENT OF THE ERROR-PRONE MODEL	
	APPENDIX F: GLOSSARY OF TERMS	

LIST OF TABLES

Table	Title	Page
2.1	Rate of Receipt of Comments by Critical Field	2.4
2.2	Comparison of Responses of Validation and Nonvalidation Applicants To Processing System Edits	2.7
2.3	Direction and Magnitude of Corrections by Critical Field	2.10
2.4	Current Payment and Eligibility Status by Critical Field of Comment Review	2.13
2.5	Rejection Rates by Income	2.17
2.6	Rejection Rates by Dependency Status and Income . . .	2.19
2.7	Characteristics of Applications Most Frequently Rejected.	2.20
2.8	Distribution of Rejected Applicants by Rejection Reason and Corresponding Comment.	2.25
2.9	Percentage of Applicants Correcting Relevant Critical Fields on the Transaction Just Subsequent to Receiving a Rejection Edit	2.28
2.10	Verification by Rejection Reason.	2.29
2.11	The Magnitude and Direction of Correction in Response to Rejection Edits in Terms of Effective SEI Change and Average Potential Payment Change. . .	2.31
2.12	Percentage of Applicants Who Remain Rejected on the Transaction Just Subsequent to Making a Correction to a Field Relevant to a Reject Reason Code	2.33
2.13	Current Eligibility and Payment Status of Ever Rejected Applicants by Project Reason	2.34
3.1	Summary of Validation Ceilings and Number of Applicants Selected	3.3
3.2	Overall Comparison of Effectiveness of Random Selection Versus Selection According to PEC.	3.7

LIST OF TABLES (Continued)

Table	Title	Page
3.3	Summary of Validation Applicant Post-Selection Corrections Behavior and Resulting SEI Change by PEC and Random Applicants	3.10
3.4	PEC with Highest Correction Rate.	3.11
3.5	PEC with Greatest and Least SEI Change.	3.12
3.6	Most Effective PEC as Determined by Composite Change Index.	3.14
3.7	Least Effective PEC as Determined by Composite Change Index.	3.14
3.8	Most and Least Consistent PEC Groups.	3.15
3.9	Post-Selection Correction Rate to Critical Fields, Average SEI Change, and Composite SEI Change by Validation Applicant Subgroup	3.17
3.10	Percentages of Validation Applicants with Post-Selection Corrections to Suspect Fields by PEC Subgroup.	3.20
3.11	Percentage of Validation Applicants Not Expected to be Paid and their Current Eligibility Status	3.23
3.12	Hypothesized Relationship Between Applicants Not on the Recipient File and Applicants Correcting without SEI Change	3.25
3.13	Estimated Dollar Savings for Currently Eligible Validation Applicants Not on the Recipient File . . .	3.26
3.14	Most and Least Effective and Refined Criteria	3.29
4.1	Pre- and Post-Selection (Eligibility) Correction Behavior of Validation and Nonvalidation.	4.3
4.2	Post-Selection Corrections, Average SEI Change and Percent of Applicants Correcting with No SEI Change - Validation and Nonvalidation Applicants. . .	4.7
4.3	Summary of Expected Payment and Current Eligibility Status of Validation and Nonvalidation Applicants . .	4.8

LIST OF TABLES (Continued)

Table	Title	Page
4.4	Comparison of the Corrections Behavior of Validation Applicants in Response Rejection Edits and their Corresponding Criteria.	4.11
4.5	Interaction Affects of Validation and Edits on Validation Applicants	4.13

LIST OF EXHIBITS

Exhibit		Page
2.1	Summary of Applicants Corrections	2.6
2.2	Current Payment and Eligibility Status: Applicants Receiving Comments vs. Applicants Not Receiving Comments.	2.12
2.3	Summary of Current Status of Applicants Ever Rejected.	2.16
5.1	Definition of Dependent Variable	5.8
5.2	Predictor Valuables	5.9
6.1	Error-Prone Model Tree Diagram.	6.3
6.2	Definition of the 37 Groups	6.6

EXECUTIVE SUMMARY AND RECOMMENDATIONS

Background

The Basic Educational Opportunity Grant (BEOG) is the largest of the student financial aid programs administered by the Office of Student Financial Assistance (OSFA). BEOG was authorized by Title IV of the Higher Education Act of 1965, and is second only to the Guaranteed Student Loan program in total student compensation. Students who receive BEOGs are also eligible for other types of State and Federal financial aid; thus, the BEOG serves as a cornerstone of aid to students who are eligible based on a formula which determines financial need. The result of this formula calculation is a student eligibility index (SEI) which, together with cost of education at the institution the student plans to attend, and the student's enrollment status (full time or part time), determines the amount of the BEOG to which the student is entitled.

One of the OSFA's management objectives is to reduce the amount of student misreporting on BEOG applications, or to increase the rate of awards based on accurate information. To meet this objective, OSFA has instituted several procedures for detecting and correcting errors on applications; namely, application processing system edits and pre-award validation of selected applicants by financial aid administrators (FAAs). The purposes of this study are to evaluate the impact of these quality assurance procedures on the correct award of BEOGs, and to develop a model for detecting error-prone applications early in their processing.

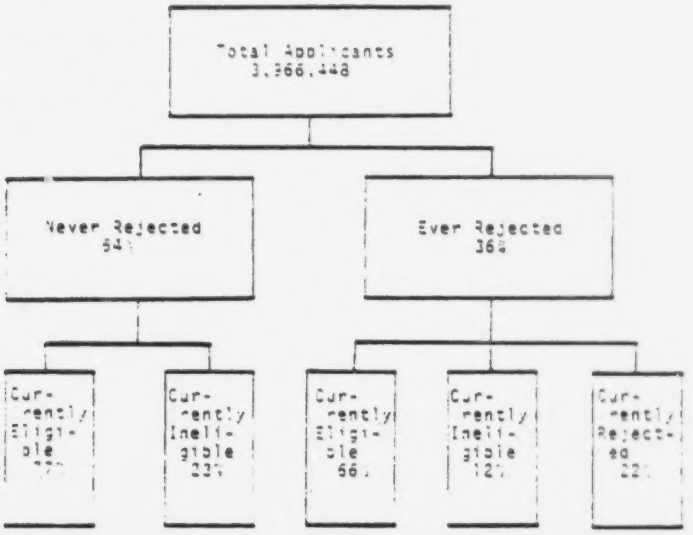
BEST COPY AVAILABLE

Description of the Application Processing Edits

Students may apply for a BEOG using one of several application forms. All applications are, however, sent to a central processing agency, where a student eligibility report (SER) results and is mailed to each applicant. If all application information appears complete and consistent, an SEI is computed. To be eligible for a grant, the value of the SEI must be 1600 or less. For eligible applicants, a maximum potential award (based on full time enrollment at the institution indicated as the applicant's first choice) is calculated. Both the SEI and award amount are printed on the SER which the student brings to the FAA for final award calculation based on actual enrollment status.

In many cases, however, the information initially reported on the application appears inconsistent or incomplete. This determination is made by the edits, which are applied to all applications as they are processed. In 1979-80, nearly half (42%) of the 3,966,448 applicants received an edit to an application item which was critical to calculation of the SEI. These edits result in comments which are printed on the SER and returned to the student for review, corrections or verification. Many of these comments cause the application to be rejected; that is, an SEI will not be computed unless the student responds to the comment by verifying or correcting the items indicated. In this academic year, 36 percent of all applicants received rejection comments. The following chart shows the rejection and eligibility status of Basic Grant applicants.

REJECTION AND ELIGIBILITY STATUS OF 1979-80 BASIC GRANT APPLICANTS



Description of the Pre-Award Institution Validation Process

In academic year 1978-79, BSFA initiated a procedure for validation of certain data items on selected applications by financial aid administrators at the institutions to which these students apply. (The chart on the following page shows the distirubtion of validation and nonvalidation on applicants for the 1979-80 academic year). BSFA had developed criteria, which according to several previous studies, indicated inaccurate reporting on applications. These preestablished criteria (PEC) have been refined and were used to select 166, 348 applicants for validation. In addition , a smaller group of applicants were randomly-seleted to be validated. This random group was selected for comparison with the group selected according to the PEC, as an ongoing check on the effectiveness of the PEC and of the validation process.

The student is informed by a comment on the SER that he/she has been selected for validation, and is instructed to bring documentation of certain information supplied on the application to the FAA. This documentation and the application are reviewed by the FAA, and the student is instructed whether to verify or correct the items in question.

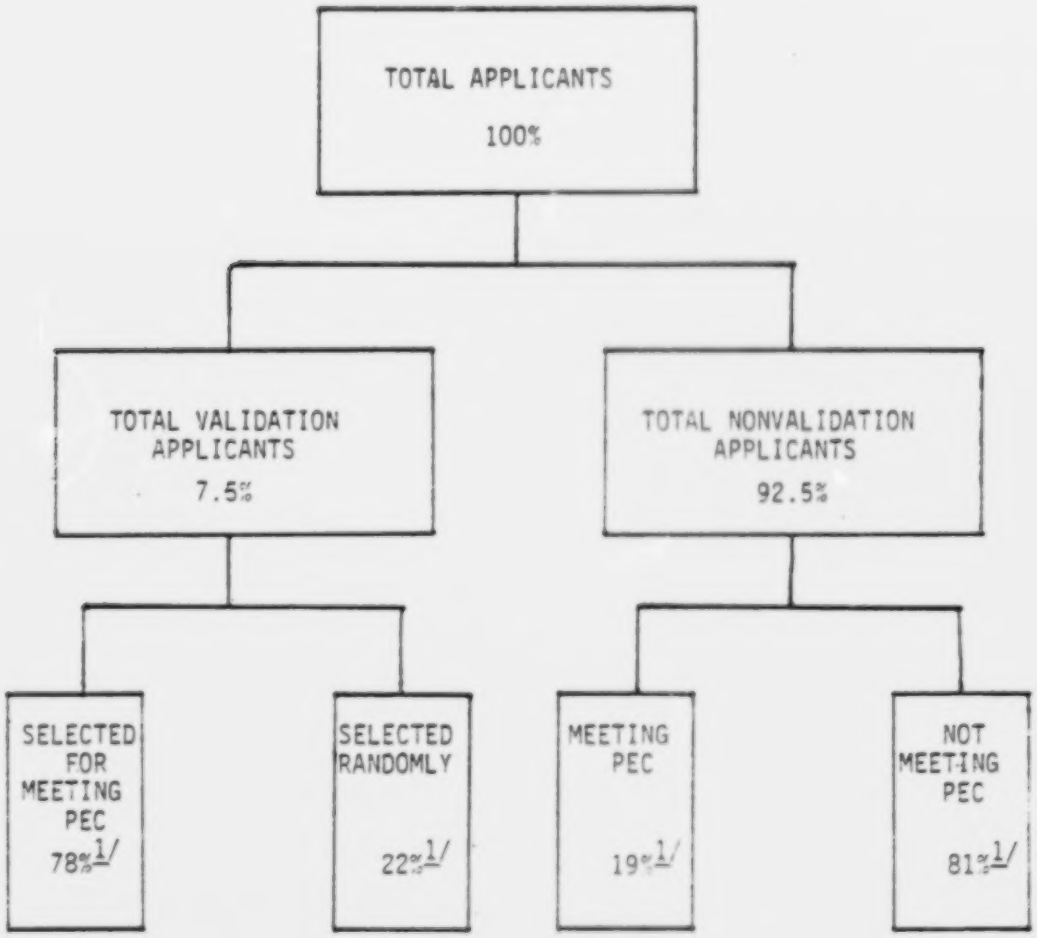
Summary

This study has two major thrusts: one, to assess the impact and effectiveness of validation and the edits; and two, to explore the viability of a statistical sequential search technique in improving upon current methods used to identify error-prone applications. The major study findings follow.

IMPACT OF PROCESSING SYSTEM EDITS (Chapter 2)

- The processing system edits had a substantial impact on the corrections behavior of applicants. About 42 percent of all applicants received at least one edit addressing a key application field. About one-half of the applicants responded to the edits by correcting the information reported originally. However, nearly half of these corrections were inconsequential in terms of award potential.

DISTRIBUTION OF VALIDATION AND NONVALIDATION APPLICANTS



1/ Percentages are based on the total validation population and the total nonvalidation population, not on the number of total applicants.

- The edits impeded many applicants from receiving a grant. The majority of eligible applicants who had never received an edit had become recipients by May, 1980. In contrast, most applicants who received an edit were either still negotiating the system or had been deterred. Since this study is limited to analysis of existing data, it is not feasible to determine the reason for the lower rate at which these applicants attained recipient status.

- About 36 percent of all applicants were prevented from receiving an eligibility determination by the edit system for incomplete or apparently inaccurate application information. As of May, 1980, 78 percent of these applicants had re-entered the system and received an eligibility determination.

- Based upon the results of institution validation, the edit system appeared to identify inaccurate and incomplete data. For validation applicants, a majority of whom corrected in response to an edit prior to selection for validation, the edits were so effective in soliciting valid information that they rendered validation unnecessary for approximately 85 percent of the cases. This finding assumes that institutional validation results in "valid" information, which is an untested assumption.

- Edits associated with the Social Security and Veteran's Benefits tape matches appear to be the most effective of all edits in identifying inaccurate application information; the edit given to applicants reporting a very low income appears to be the least effective in identifying invalid data.

VALIDATION (Chapters 3 and 4)

- Selecting applicants for validation based on the preestablished criteria is in general, a more effective approach to selecting applicants for validation than the random selection process. A greater percentage of PEC applicants than random applicants correct post-selection. These corrections result in a larger average positive SEI change, hence greater savings to the Department of Education.

- PEC groups A and D - groups based not on the applicants past corrections history but instead on information on the applicants current financial status - are the best predictors of misreporting. Groups B and C - groups based on correction history - are no more effective than random selection in identifying applicants likely to correct post-selection.

- There is great variation in the efficacy of the PEC subgroups. Seven of the PEC subgroups perform exceptionally well in identifying misreporters (A-5, A-6, A-10, A-11, A-14, A-15 and B-3). These subgroups elicit high numbers of applicants correcting, have relatively large SEI changes, and are discriminating in selecting only those applicants whose corrections will affect their potential award.

- PEC subgroups A-1 (household size or marital status missing) and (low income applicants) A-4 continue to be most ineffective in identifying potential misreporters. Data indicate that a low percentage of small income applicants correct in response to validation and that of the percent correcting, only a very few applicants make correction that affects their SEI. These applicants remain eligible more often than other groups of applicants, yet do not appear on the recipient file as often, thus suggesting that the current edits and validation system may be unnecessarily hindering them in obtaining an award. The fact that financial aid administrators may find it difficult to validate income sources or the lack of sources particular to this subgroup, may mean that, although the data indicate this group to be accurate reporters, in reality the applicants may be submitting inaccurate information that is not being detected by current validation procedures.

ERROR PRONE MODEL (Chapters 5-8)

The application of a sequential search technique to the randomly selected validation sample has shown to be promising in identifying error-prone applicants and segregating them into identifiable groups. For the purpose of this analysis, "error-prone" applicants are defined as those who make corrections resulting in an eligibility index change greater than 50 points after being selected for validation. The results of this effort follow.

- Thirty-seven groups have been identified. They differ from each other in the proportion of applicants who:
  - do not (seriously) misreport
  - misreport to their advantage
  - misreport to their disadvantage
  - fail to re-enter the system after selection for validation
- Within the 37 groups, eight groups account for 28.7 percent of all misreporting to the applicant's advantage. Applicants belonging to these groups make up eleven percent of the eligible population.
- In contrast, the pre-established criteria currently used by OSFA to select applicants for validation account for 5.9 percent of the eligibles and locates only 9.1 percent of the applicants who misreport to their advantage.
- Validating applicants in these eight groups would be far superior to randomly selecting applicants for validation and selecting applicants according to the pre-established criteria. A total of 45.5 percent of the applicants in these groups increased their eligibility index by 50 or more points after

validation. Only 20.8 percent of the randonly selected validation applicants and 29.2 percent of the applicants selected according to the pre-established crieteria made this type of change.

Recommendations

Based upon these study findings, the following recommendations are offered to OSFA:

RECOMMENDATION #1: Conduct a field study to determine whether institutions conduct validation.

The major findings in this report concerning impact are based on the assumption that institutions conduct validation according to OSFA - prescribed procedures and that validation results in correct information. This is an untested assumption, and a field study should be conducted to determine whether it is a valid assumption.

RECOMMENDATION #2: Conduct a field study to determine why corrections are made.

The current system requires that other key assumptions be made, in particular that corrections made after receipt of an edit or after selection for validation are being made in response to the systen. The degree to which system actions cause the corrections versus corrections being made because applicants are confused or are trying to test the system or for some other reason are unknown. A field study needs to be conducted to determine the extent to which corrections mean what they are assumed to mean.

RECOMMENDATION #3: Conduct a field study to find out why applicants drop out of the system after receiving an edit or validation.

Applicants who receive an edit or are selected for validation and who subsequently fail to re-enter the systen or to obtain an award after negotiating the system are a mystery. It is hypothesized that some applicants do not return because they are misreporting and have been found out, while others are frustrated and have given up. In order for OSFA to make appropirate decisions about the edits and validation, a field study needs to be done to find out why these applicants drop out of the system.

RECOMMENDATION #4: Improve timeiness in obtaining award data.

The dollar impact of the edits and validation is key in providing a basis for sound decisionmaking. However, there is a significant time lag between when the student receives an award and when OSFA obtains informtion on the award amount. OSFA needs to review its financial system and make improvements to obtain timely information on award expenditures.

Within the context of the current system and the data which were available for this study, the following recommendations are made:

RECOMMENDATION #5: Unless OSFA alters the current edit and validation system, the edits which focus on low-income applicants should be reviewed and possibly eliminated. Study findings have shown that these edits tend to cause more applicants to drop out of the system than do other edits. Further, those applicants who receive the edits and do re-enter the system show virtually no impact to corrections.

RECOMMENDATION #6: Continue to expand and refine the tape match edits. Study findings have shown that the matches with SSA and VA are more effective in identifying and correcting misreporting than other edits. OSFA should refine its current efforts by studying ways to improve the current match rate with SSA and VA. Given the effectiveness of these efforts, OSFA should try to clear legislative channels to allow additional tape matches with other data bases, particularly the Internal Revenue Service.

RECOMMENDATION #7: With the exception of the low income edits, continue the current system. Based upon validation findings, it appears that most edits are effective and eliminate the need for validation for many applicants.

RECOMMENDATION #8: Revise the validation selection procedures to incorporate the findings of the error-prone model (EPM). Preliminary findings show that several of the groups identified in the EPM are superior to the current PEC in identifying misreporting, and that they should be used with the successful PEC subgroups for identifying error-prone applications. Before using the results of the EPM, it is important that it be updated using more recent applicant and payment data than were used in this exploratory study.

RECOMMENDATION #9: The EPM identified the estimation of income and taxes by applicants who submit applications before filing a Federal Income Tax Return as a major source of error. OSFA may eliminate such errors through a range of actions with the following as extremes:

- Not accepting applications prior to the Federal tax filing deadline

or

- Requiring copies of tax returns from all applicants who apply before they have completed the Federal tax return.

Given that the first alternative would place a serious burden on institutional aid packaging schedules, which in turn would affect student choice of school, it is recommended that OSFA entertain the second alternative. This could be accomplished through several ways, the most simple of which would be to either modify processing procedures and require the tax return with the application or modify institutional validation procedures to include verification of tax return data only for these gorups.

RECOMMENDATION #10: OSFA should identify applicants who are misreporting to their disadvantage, and develop a technical assistance program to facilitate their receipt of the correct

entitlement. The EPM has identified some groups of students who error to their disadvantage. These students should receive edits advising them to seek assistance in completing their application, either from their financial aid administrators or from one of OSFA's information service contractors. Applicants should be advised that they may be eligible for a larger award to encourage action on their part. The Student Eligibility Report should contain a code to cue the technical assistance provider to possible nature of the problem.

1

STUDY OVERVIEW

1.1 Purpose and Scope

For the 1978-79 academic year, the Bureau of Student Financial Assistance (BSFA) introduced two major program initiatives intended to ensure that applicants were submitting accurate information on their Basic Grant application form. The new program procedures tightened certain edits in the processing system, and required financial aid administrators from post-secondary educational institutions to validate certain application information of selected students. Last year, the BSFA introduced new comments into the edit system and expanded the pre-established criteria (PEC) used to select students for validation. Part I of this report examines the impact of these changes on the correct award of Basic Grants during the 1979-80 academic year.

Part I of the report is divided into discussions of three major issues: the impact of the current processing system edits, the effectiveness of the pre-established criteria as indicators of misreporting, and the impact of validation on the application information. Chapters 3, and 4 of this part provide indepth analysis of these factors. The remainder of this chapter discusses the research objectives and questions, study design and population, and methodology and data caveats associated with each of these issues.

1.2 Key Study Objectives Questions

The broad objectives of the Edits and Validation analyses presented in this Part are three-fold. They are:

- to assess the impact of the processing system edits on applicant behavior;
- to assess the adequacy of the pre-established criteria as effective indicators of misreporting; and
- to determine the impact of the validation effort on the correct award of Basic Grants.

The specific study objectives and study question which guided these analyses are divided into three sections which correspond to the broad study objectives listed above.

- The Processing System Edits
  - -- to determine the type and extent of corrections made by applicants in response to the edits, and whether their corrections behavior differs substantially from that of applicants who were not screened by the edits;
  - -- to determine who is being rejected by the processing system and, in particular, whether the edits are unnecessarily rejecting low income, needy applicants;
  - -- to determine the extent to which the rejection edits are impeding applicants from re-entering the system and receiving an eligiblity determination;
  - -- to determine who is re-entering the system following a rejection; and
  - -- to determine the relative efficacy of the individual rejection edits in identifying incomplete and inaccurate applications.
- The Pre-established Criteria
  - -- to determine whether the pre-established criteria are more effective than the random selection process in identifying applicants likely to make post-selection corrections resulting in significant SEI change;
  - -- to determine the relative efficacy of the four pre-established criteria groups (A, B, C & D) and subgroups in identifying students reporting incorrect application information;
  - -- to determine the consistency with which the PEC criteria identify misreporters: that is, the degree to which the PEC identify only those applicants whose corrections result in SEI change;
  - -- to determine the relationship between the reason a PEC applicant was selected for validation (the PEC met) and the critical fields corrected post-selection by the applicant;

- -- to determine whether PEC applicants have greater difficulty in achieving recipient status than random applicants; and
- -- to determine the potential dollar saving resulting from the use of the PEC as selectors for validation.

- Validation (and the Interaction of Validation and Edits)
  - -- to determine the extent and direction of corrections made by validation applicants prior to their selection for validation and whether their correction behavior differs substantially from that of applicants not selected for validation;
  - -- to determine the extent and direction of corrections made by validation applicants post-selection and whether their correction behavior differs significantly from nonvalidation applicants;
  - -- to determine whether validation affects the consistency with which post-selection corrections resulting in SEI change are made: that is, whether a higher or lower percentage of validation applicants make corrections that do not affect their potential award;
  - -- to determine whether the percent of applicants on the recipient file is similar for the validation and nonvalidation groups;
  - -- to determine the effect of validation on the size of the actual Basic Grant award; and
  - -- to determine the relative impact of validation and the edits on applicant corrections behavior.

1.3 Study Design

The three major issues discussed in this Part required slightly different study designs for analyses. The following summarizes the study design for each issue.

a) The Impact of the Current Processing System Edits

In general, the purpose of the analyses of the processing system edits is to examine the edit's impact on the following areas:

- the frequency and type of corrections to key applicationitems; and
- the ability of applicants to swiftly meet the demands of the processing system and receive an eligibility determination.

The analysis of the edits has been divided into two sections. The first section examines the global impact of the edits. The corrections behavior and current eligibility (rejected, eligible, ineligible) and payment (expected to be paid, not expected to be paid) statuses of applicants who received edits are compared with those who did not. In addition, the frequency, magnitude and direction of corrections in response to the edits are compared with the voluntary, student-initiated corrections.

The second section assesses the relative efficacy of the individual rejection and assumption edits. Each rejection edit is assessed by examining applicant corrections and the ability of applicants to re-enter the system following rejection. Two major assumptions underlie the examination of the adequacy of the rejection edits:

- If an edit solicits frequent corrections that have more than a negligible impact on the applicant's SEI then it is successfully identifying incomplete or inaccurate application data.
- In order to be worthwhile, a high percentage of applicants must re-enter the processing system following the receipt of the edit.

b) The Adequacy of the Pre-established Criteria

To determine the adequacy of the pre-established criteria as indicators of misreporting, a sample of individuals selected for validation because they met one or more PEC criterion was compared with a sample selected randomly. These groups were then compared on three levels. First, the large group of all PEC applicants was compared to the group of random applicants to determine whether the concept that certain misreporters can be identified by patterns of applicant behavior or information reported by the applicant is viable. Next, the PEC groups (A, B, C and D) were compared to each other to ascertain the relative effectiveness of each group in detecting misreporters. Finally, each PEC criterion or subgroup was analyzed and its performance in identifying applicants submitting incorrect information contrasted with the performance of the other subgroups.

Each of the PEC groups and PEC criterion was analyzed according to several measures of effectiveness. The primary measures used for analyses were:

- the percentage of applicants correcting any field post-selection and those correcting only critical fields; and
- the magnitude of the resulting average SEI change[1/].

The use of these measures is based on the assumption that corrections made post-selection, and hence after the student has been validated, reflect instances where validation detected inaccurate information on the application and the student corrected. It is assumed that the correction reflects accurate information. Theoretically at least, the greater the percentage of applicants correcting post-selection, the higher the number of misreporters identified. The magnitude of the SEI change which results from the correction should also reflect the disparity between the old, incorrect application information and the new, correct information. Therefore, the more inaccurate the initial information or the more serious the misreporter, the greater the SEI change resulting from the correction.[2/]

[1/] The average SEI change is derived by summing the SEI change scores for applicants raising and lowering their SEI and dividing this figure by the number of applicants correcting, including those whose corrections did not result in SEI change.

[2/] We realize that these assumptions do not always hold true. For example, it is plausible that an applicant's initial information was correct and that validation elicited an inappropriate change. It is also possible that if the initial data were accurate, the correction reflects a legitimate change in the applicant's status. Furthermore, even when the initial information is incorrect, we have no way of verifying concretely that the new information is correct and that the applicant is not substituting new falacious information for old. Since we also have no guarantee that validation actually took place, or if did take place, that the students' documents were checked thoroughly, we can only assume that post-selection correction are indicative of misreporting.

In addition, several other measures were used to assess effectiveness. Although these measures are somewhat secondary to the above, they allow for finer analyses and interpretation of the data. These measures are as follows:

- the percentage of applicants correcting post-selection whose corrections do not result in SEI change;
- the percentage of students raising and lowering their SEI and the amount raised and lowered;
- the relationship between the reason an applicant was selected for validation and the fields the applicant corrected post-selection;
- the rate at which PEC and randomly selected applicants enter the recipient file, and the current eligibility status of the applicants; and
- the average difference in expected award between the transaction on which the student was selected for validation and the current transaction.

c) The Impact of Validation and the Interaction of the Edits and Validation

The third issue of concern in Part I of this report is the impact of validation on the correct award of Basic Grants. For this Part, we compared a sample of validation applicants (both applicants meeting PEC and those randomly-selected) with a similar sample of nonvalidation applicants. These groups were compared first on their pre-selection correction behavior, (including the percent correcting and the fields changed) to determine whether the two groups were initially similar enough to allow comparisons. The post-selection correction behavior and the resulting average SEI change for both groups was then compared to assess the degree of difference between the two groups and the impact of validation. Finally, payment data was analyzed to provide some indication of the actual dollar savings that might result from validation.

The assumptions underlying these analyses are the same as those for the analyses of the effectiveness of the pre-established criteria: that is, post-selection corrections reflect the changing of inaccurate data to accurate data and thus reflect misreporting; and the magnitude of the average SEI change is indicative of the degree of misreporting.

Although not considered for analytical purposes a separate issue, we also attempted to determine the relative impact of the edits and validation by developing tables which compare the validity of data fields at various points in the application processing and validation cycle. This allowed us to determine the approximate degree to which validation is necessary to insure accurate information.

1.4 Study Population

The population studied for these analyses consisted of 3,966,448 applicants, or all individuals on the applicant and recipient files as of May 1980, except for a fairly small group of independent students whose data history could not be incorporated properly into the files. This was due to problems in their corrections history resulting from a Congressionally mandated recomputation of their SEI (see the methodology section of this chapter for further discussion).

To facilitate timeliness and to minimize cost, a ten percent (10%) sample of this entire 3.9 million applicant data base was used for many of the analyses. The sample drawn consists of one of every two validation applicants, and one of every fourteen other applicants. (The PEC validation applicants were oversampled in relation to the other groups to ensure adequate numbers of subjects in each criterion group.) Population samples were then produced by assigning the appropriate weight to the sampled groups. The reliability of this sample approaches 100 percent due to the large sample size.

Different subsets of the sample were used for various analyses. Therefore, the total number of applicants reported in various tables of this report differ. In general, the edits chapter of this part of the report is based on analyses of the total number of applicants ever receiving comments, and the total number ever rejected. The population totals for these groups follow:

PRIMARY GROUPS SAMPLED FOR EDITS ANALYSES

Total Number of Applicants Receiving Comments	Total Number of Applicants Ever Rejected
1,650,522	1,423,540

The only sample population used in analyzing the adequacy of the PEC subcriterion is the set of all validation applicants. This sample is divided into those applicants meeting the pre-established criteria those selected randomly.[1/] The file reports the following distribution of validation applicants.

SAMPLE POPULATION FOR PEC ANALYSES

Total Number of Validation Applicants	212,362
Number Meeting PEC	166,348
Number Randomly Selected	46,014

Finally, the validation chapter of this Part is based on a sample of all validation and nonvalidation applicants. All applicants included in this sample have at one point in time been eligible for an award. The sample is divided according to whether the applicants are selected for validation and according to whether the applicants meet the PEC, or were selected randomly (or did not meet PEC).

1/The "randomly" selected group was chosen by flagging every nth applicant. It is not truly random, however, because the selection ceilings were occasionally changed by OE, and because those who submitted corrections had multiple transactions (and thus a greater chance of selection). All applicants, therefore, did not have an equal probability of selection.

SAMPLE POPULATION FOR VALIDATION ANALYSES

	Total Applicants	Column %	Validation Applicants	Column %	Nonvalidation Applicants	Column %
Meeting PEC	673,260	(24)	166,348	(78)	506,912	(19)
Not Meeting PEC	2,162,926	(76)	46,014	(22)	2,116,912	(81)
TOTAL	2,836,186[1]/	(100)	212,362	(100)	2,623,824	(100)

1.5 Study Methodology

The analyses of the current edits, application processing and validation system are based on bivariate contingency tables and frequency distributions designed and developed specifically for this project. The tables analyzed were run on a file which contained merged secondary data provided by the Basic Grant central processor and the Department of Education.

It should be mentioned that before any tables were produced a subsample of the total population was excluded from the data file. This was necessary because during the 1979-80 academic year Congress approved a new formula which changed the way financial need was computed for independent students. The new formula was put into effect on May 9, 1979 and consequently the SEI's of all independent students applying before that date had to be recomputed. The recomputation changed the SEI's of a significant number of independent applicants whose initial SEI had been greater than zero. However, since the new formula was intended to increase aid to independent students, those applicants whose initial, pre-May 9, SEI was zero did not have an SEI change.

1/As mentioned in the text, all applicants included in the sample used for the validation analyses have been eligible for a grant at some point in their application history. This narrows down the population from all Basic Grant applicants and explains the difference in the total number of applicants reported on page I.8 and this page.

Regardless of whether a student's SEI changed as a result of the recomputation, a new SER was generated all independent applicants who filed prior to May 9. The new SER's were only mailed to those students whose SEI had changed. The newly generated SER's, however, were recorded as systems generated transactions by the central processor and showed up as correction in our data file.

In order to retain as many independents affected by the formula change as possible, yet not confound the data, we ignored the systems generated transaction for students whose initial pre-May 9 SEI was zero (and thus did not change), and excluded all other affected students from the population. Applicants, whose SEI changed as a result of the recompute, could not be included as part of the sample population because of the inability to distinguish between SEI changes caused by the recompute and changes caused by other factors such as misreporting.

1.6 Study Caveats

As mentioned previously in this chapter, there are several minor caveats that should be considered in reading and using these reports. These caveats can be divided into two categories: those relating to the assumptions underlying the use of certain measurement variables, and those related to data processing concerns.

Most of the caveats related to the assumptions underlying the use of certain measurement variables have been discussed earlier, so they will only be summarized here. The first caveat concerns the assumption that corrections are indicative of misreporting. At this time we do not know what percent of all corrections are due to applicants correcting inaccurate information. Although most corrections, particularly post-selection validation corrections, would seem to be indicative of misreporting, some corrections must reflect changes made for other reasons. The Internal Revenue Study currently in progress, will help determine the degree to which corrections reflect changes from incorrect to correct information. In addition we have no clear evidence whether students are accurately validated, or if validated, how thoroughly their

documentation is checked; we only know that applicants were selected for validation. Before making conclusive statements regarding the impact of validation, per se, it will be necessary to determine to a better degree the extent and the quality of the validation process.

There are three other constraints imposed primarily because of data processing reasons that must be considered in using this report. First, the data is not representative of the total Base Grant population. As mentioned in the methodology section above, a small, but significant percent of the population i.e., (independency students whose SEI was increased as a result of the SEI formula change) was excluded from the data sample. This means than no inferences can be drawn regarding the behavior of this group of applicants. Second, it was necessary to compute an artificial student eligibility index for rejected applicants in order to conduct the edits analyses. Rejected applicants do not have SEI and therefore, the extent to which our assumptions are correct in determining the artifical SEI, affects the accuracy of the final data. Finally, the reader should be aware that the central processor's failure to flag nonvalidation applicants meeting PEC groups B, C, & D, negates comparison of scores of the total validation PEC group and the total nonvalidation PEC group. Only the scores of the A group and subgroups, and the random selection (not meeting PEC) groups can be compared. Otherwise, the exclusion of PEC nonvalidation B, C, & D applicants will result misleading averages.

2

IMPACT OF THE PROCESSING SYSTEM EDITS AND COMMENTS

The Basic Grant application processing system includes several features that were designed to minimize the number of student eligibility determinations made on the basis of invalid, inaccurate or incomplete data. One feature, discussed in detail in Chapters 3 and 4, is the selection of applications for validation. Another feature is a series of computerized edits which check for missing information and the logic and consistency of all application data provided.

If an application triggers one of these processing system edits, a message is printed on the Student Eligibility Report (SER) advising the student to review the application and take further action, if necessary. Under the most restrictive conditions, the edit comment indicates that the application has been rejected and that the student must provide missing information or verify or correct existing data before an eligibility determination can be made. In other cases, the comment indicates that the processing system, in calculating eligibility, assumed a value for a missing or apparently inaccurate application item based on other provided information. In the third case, the comment serves as merely informational, or as an attention-attracting device warning of questionable data.

The purpose of this Chapter is to examine the impact of the processing system edits and comments on: 1) the frequency, magnitude and direction of applicants' corrections to key application items; and, 2) the ability of applicants to expeditiously pass through the processing system and obtain an eligibility determination. Section 2.1 gives a description of the edits and comments and an overview of the impact of

the edits on applicant behavior. In this section, the corrections behavior, the current eligibility status (rejected, eligible, ineligible) and payment status (expected to be paid, not expected to be paid) of applicants who received edits are compared with those who did not. Section 2.2 provides a more detailed analysis of the effect of the most restrictive type of edit, the rejection edit. The first part of this section examines the income and household characteristics of applicants most often rejected and the second part assesses the effectiveness of the individual rejection comments and edits in terms of applicant corrections response.

2.1: Global Impact of the Edits

2.1.1 Description and Distribution of Comments

There are 317 different computerized comments generated by the processing system. For this analysis, only the 176 comments that pertain to applicants filing a regular application and which address key application items were examined. (See Appendix A for a list of the comments used in this study). Key application items are defined as those which have a major influence in the computation of the Student Eligibility Index (SEI) and/or affect whether an applicant is selected for validation. A list of these key application items follows:

- Adjusted Gross Income (AGI)
- Taxes Paid (TP)
- Portions Earned
- Nontaxable Income (Social Security Benefits, Other Nontaxable), (NTI)
- Veteran's Educational Benefits (VEB)
- Net Assets (NA)
- Applicant's Resources (AR)
- Unusual Expenses (medical/dental, casualty, fire/theft), (UE)
- Household Size (HS)
- Dependency Status (Model)
- Unreimbursed Elementary and Secondary School Tuition (UT)
- Post High School Enrollment (PHE)
- Marital Status (MS)

- Tax Filing Status (tax return figures are: (a) from completed return, (b) estimated, (c) not appropriate because applicant will not file a return), (TFS)

Hereafter, these key data fields are referred to as "critical" fields. The abbreviations noted above are used throughout the report.

As of May, 1980, the 176 comments examined in this study had been generated 4,164,796 times; 1,650,522, or 42 percent of all applicants, had received at least one of these comments, for an average of 2.5 comments per applicant. Since many applicants receive rejection edits prior to being selected for validation, it is not surprising that validation applicants received critical field comments at a higher rate than nonvalidation applicants: 84 percent of validation applicants received critical field comments, for an average of 4.7 per applicant, while only 39 percent of all applicants not selected for validation received comments, for an average of 1.3 per applicant.

As mentioned earlier, there are three types of edits and comments: rejection, assumption and informational. Many applicants received more than one type of comment: about 65 percent received rejection comments, 22 percent received assumption comments and 56 percent received informational comments not related to the rejection or assumption edits.

There was considerable variation by critical field in the number of comments generated. The following table lists the number of applicants receiving comments by the critical field which the comment addressed. (A list of the number of applicants who received each individual rejection comment can be found in Section 2.2.)

Table 2.1 illustrates that the largest percentages of applicants received comments to items most critical in determining eligibility. This may have occurred because the processing system has the most stringent requirements for these fields; it is also indicative of questionable data being reported (or not reported) in these fields. A relatively large proportion of applicants received comments addressing post high enrollment, a key application item in determining eligibility. By comparison, relatively few received comments to household size, an equally important field in the award computation.

TABLE 2.1: RATE OF RECEIPT OF COMMENTS BY CRITICAL FIELD

CRITICAL FIELD	NUMBER OF APPLICANTS RECEIVING COMMENTS 1/	PERCENT OF TOTAL WHO RECEIVED COMMENTS 2/
AGI	1,009,936	61%
TP	887,825	54
PORTIONS	709,186	43
NTI	623,604	38
PHE	388,122	24
UE	214,828	13
TFS	210,828	13
NA	191,796	12
HS	179,116	11
VEB	143,998	9
MS	113,316	7
AR	73,022	4
UT	68,908	4
MODEL	21,384	1
TOTAL APPLICANTS	1,650,522	100

1/ Most comments address more than one field. Consequently, the frequencies in this table are duplicated between fields and do not add up to the unduplicated total number of applicants who received critical field comments. For example, an applicant who received a comment which addresses four different fields would appear in four of the above frequencies.

2/ Percentages are based on the unduplicated total number of applicants who received critical field comments.

Since the edits to both household size and post high enrollment are equally stringent, this table suggests that a much higher percentage of applicants were apparently misreporting post high enrollment or leaving the field blank.

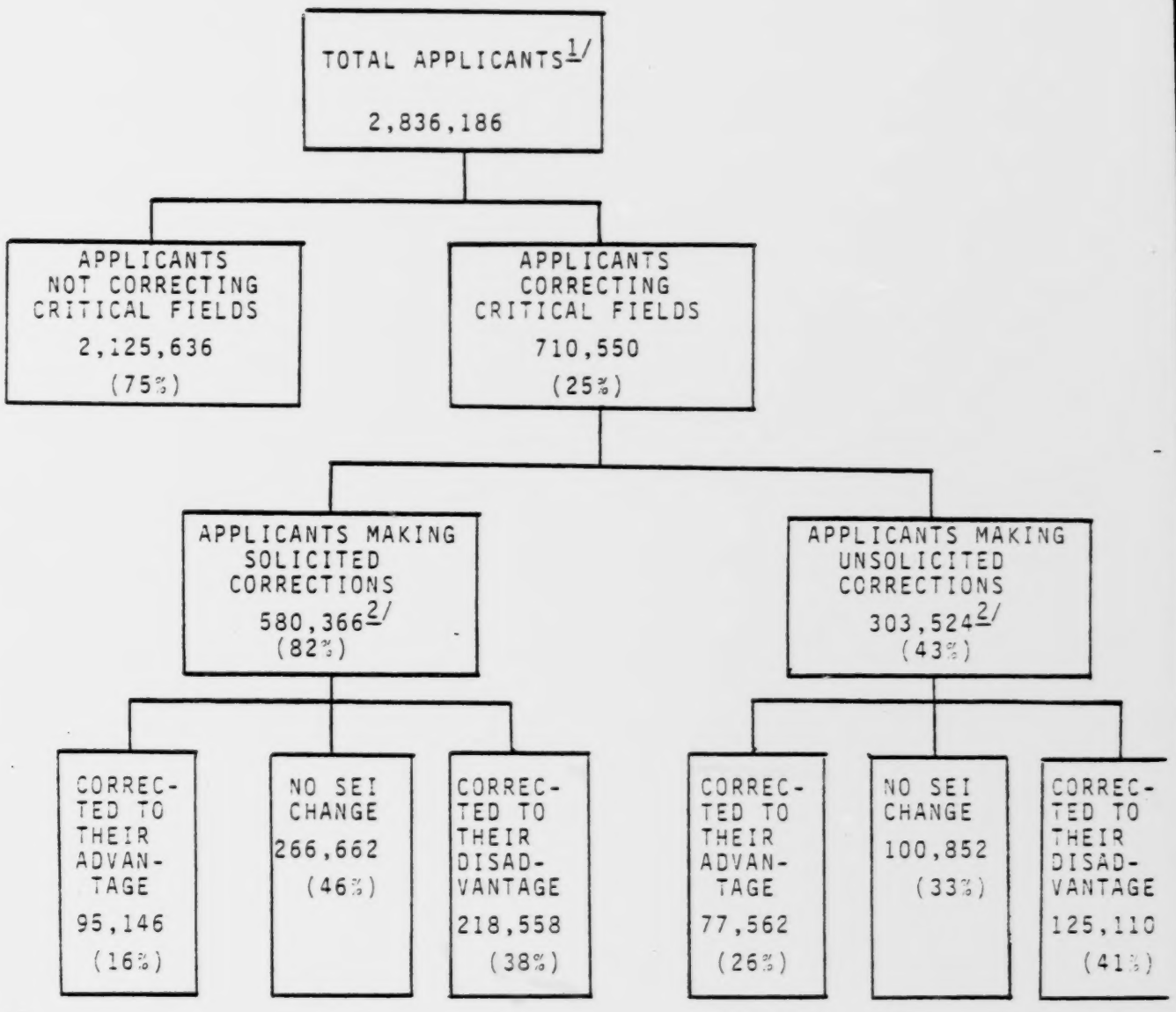
2.1.2 Corrections In Response to Comments

An applicant has the option of changing any application information once the SER has been received. A correction may be made at the applicant's initiative (instructions on the back of the SER explain how to make a correction), it may be made at the direction of the financial aid administrator, or it may be made in response to a processing system edit. For this study, all changes to critical fields which were in response to a comment are referred to as "solicited" corrections; all other corrections are considered "unsolicited". Exhibit 2.1 presents a summary of application corrections behavior.

As the exhibit indicates, about one-fourth of all applicants who were ever determined to be eligible corrected a key application field. The majority who corrected did so in response to a processing system edit. Nearly one half of all solicited corrections were inconsequential; that is, they had no effect on the applicants' SEI. The amount of insignificant responses that were elicited suggests that certain individual edits may not be cost-effective. By comparison, the SEI remained the same for only one-third of the applicants making unsolicited corrections. Applicants making unsolicited changes corrected more frequently to their advantage and to their disadvantage than applicants who responded to the edits. Regardless of whether corrections are solicited or unsolicited, those that have an impact have a greater tendency to decrease, rather than increase, award potential.

To determine further the overall impact of the edits, and to compare their influence on applicant corrections behavior with that of pre-award validation, the response modes of nonvalidation and validation applicants were compared. Table 2.2, which follows, summarizes the frequency and magnitude of solicited and unsolicited corrections made by validation and nonvalidation applicants both pre- and post-selection/eligibility. As a point of reference, post-selection/eligibility corrections occur after the transaction when the applicant is selected for validation or after the transaction when an initial eligibility determination has been made; pre-selection/eligibility corrections occur before the applicant has been

EXHIBIT 2.1: SUMMARY OF APPLICANT CORRECTIONS



1/ Refers to the total number of applicants who ever received an eligible SEI.

2/ Note that many applicants made both solicited and unsolicited corrections.

TABLE 2.2: COMPARISON OF RESPONSES OF VALIDATION AND NONVALIDATION APPLICANTS TO PROCESSING SYSTEM EDITS

		PRE-SELECTION/ELIGIBILITY				POST-SELECTION/ELIGIBILITY			
	Total Number of Applicants	% Correcting Critical Fields In Response To Edits	Average Effective SEI 2/ Change	% Making Unsolicited Corrections To Critical Fields	Average Effective SEI Change	% Correcting Critical Fields In Response To Edits	Average Effective SEI Change	% Making Unsolicited Corrections To Critical Fields	Average Effective SEI Change
Validation - Total	212,362	53%1/	+33	23%	-278	11%	+138	26%	+139
Validation - PEC	166,348	63	+30	27	-307	13	+158	28	+156
Validation - Random	46,014	20	+66	10	+5	5	-32	21	+54
Nonvalidation - Total	2,623,824	17	+70	7	+71	1	+4	3	+62
Nonvalidation - PEC	506,912	69	+99	21	+127	3	+15	4	+57
Nonvalidation - Random	2,116,912	4	-39	3	-20	1	-3	3	+64

1/ The four percentages in each row do not total 100 percent because of duplication. A validation applicant, for example, who makes one pre-selection correction and one post-selection correction is counted twice. Also, not all applicants make corrections, in particular those not selected for validation.

2/ Average effective SEI change is calculated from the time of correction to the most recent transaction. Since other corrections due to different impetuses may be included from correction to most recent transaction, there is a possible misattribution of effect.

selected or given an initial eligibility determination. The average overall effective SEI change is calculated only for those correcting, from the time of correction to the most recent transaction. Effective change refers to the actual point change if eligibility indexes above 1600 (the maximum eligible SEI) are set at 1600. For example, an absolute change for an applicant whose SEI goes from 1000 to 2200 is 1200; the effective change is 600, or 1600 minus 1000.

As Table 2.2 indicates, most corrections to critical fields occur prior to selection for validation or the determination of eligibility. The impact of the edits on SEI is also greatest during this time. The majority of pre-selection/eligibility corrections were solicited by comments whereas more post-selection/eligibility changes were unsolicited.

As mentioned earlier, validation applicants receive comments at a greater rate than nonvalidation applicants. They also respond to the edits more frequently: over half of all validation applicants corrected in response to the edits prior to selection while fewer than 20 percent of nonvalidation applicants made pre-eligibility changes. The data suggest that the edits have a greater impact on the validation applicant's corrections behavior than institution validation. Validation applicants made notably fewer unsolicited changes after selection (the assumption is that these changes are in response to institution validation) than pre-selection corrections in response to the edits. Furthermore, solicited post-selection corrections, although less frequent, had an equal effect on applicants' SEI as unsolicited post-selection corrections. (See Chapter 4 for a detailed examination of the interaction of effects of validation and the edits on validation applicants.)

In general, both the validation and nonvalidation applicants who met the pre-established criteria (PEC) made more frequent solicited corrections that resulted in larger SEI increases than applicants who did not meet the PEC. This gives preliminary indication that the rejection edits which correspond with the PEC might be working effectively.

Table 2.3, which follows, presents the corrections behavior in response to the edits by the critical field corrected. Once again, the population has been divided into nonvalidation and validation applicants. There were no notable changes in the fields most frequently corrected by each applicant group pre- or post-selection/eligibility. The rate of corrections by field can be summarized by the following three groups:

Frequently Corrected	Occasionally Corrected	Rarely Corrected
AGI	TFS	MS
NTI	HS	VEB
TP	UE	MODEL
PORTIONS	NA	AR
	PHE	UT

As previously mentioned, the majority of corrections, particularly pre-selection/eligibility, are in response to a processing system edit. However, there is considerable variation by critical field in the rate at which applicants make solicited and unsolicited corrections.

More than 85 percent of the changes to AGI, TP, and VEB followed the receipt of an edit, whereas fewer than one half the corrections to NA, HS, AR, MS, and Model were solicited. HS, AR, MS, and Model are a few of the fields which are supposed to be updated to reflect a change in circumstances following the original application submission. This may explain why a comparatively large number of applicants made unsolicited corrections to these fields.

Nonvalidation applicants, in general, corrected all fields to their disadvantage pre-eligibility, with the exception of HS and MS. This, together with the fact that fewer than 40 percent of the corrections to HS and MS were in response to comments, suggests that new or more restrictive edits might be needed for these two fields. The adequacy of the individual edits which address HS and MS is discussed in detail in Section 2.2.

TABLE 2.3: DIRECTION AND MAGNITUDE OF CORRECTIONS BY CRITICAL FIELD

FIELDS	VALIDATION APPLICANTS' CORRECTIONS						NONVALIDATION APPLICANTS' CORRECTIONS					
	PRE-SELECTION			POST-SELECTION			PRE-ELIGIBILITY			POST-ELIGIBILITY		
	Number Correcting	Average SFI Change	% In Response to Comment	Number Correcting	Average SFI Change	% In Response to Comment	Number Correcting	Average SFI Change	% In Response to Comment	Number Correcting	Average SFI Change	% In Response to Comment
Adjusted Gross Income	82,000[1/]	+73[2/]	90.0%	37,902	+100	28.8%	293,900	+105	87.6%	50,330	+11	23.0%
Taxes Paid	51,868	+100	89.8	24,568	+166	12.9	190,806	+102	88.3	36,834	+79	8.5
Portions	53,200	+16	79.7	22,506	+144	15.4	171,934	+48	66.8	33,614	+12	14.6
Nontaxable Income	43,498	-168	56.7	30,150	+85	54.4	144,536	+103	67.9	43,722	+67	22.6
Veteran's Ed. Benefits	6,482	-206	93.0	2,510	+95	14.7	26,796	+26	85.4	9,058	-47	82.1
Household Size	14,442	-309	42.2	11,532	+118	8.5	66,402	-19	36.5	18,200	+46	20.8
Post-High Ed.	10,034	-222	51.1	8,812	+120	6.7	50,176	+82	29.7	11,606	-2	28.0
Net Assets	20,418	-341	38.0	5,722	+162	20.1	96,620	+22	56.5	17,752	+83	16.5
Unusual Expenses	38,236	+275	73.3	10,816	+321	63.6	65,940	+245	49.0	6,364	+115	38.6
Applicant's Resources	5,908	-501	69.1	1,304	+95	32.5	15,404	+95	13.6	3,700	+67	15.6
Marital Status	8,238	-269	44.2	2,564	+167	10.0	50,512	-17	37.4	8,600	+92	20.6
Unreimbursed Tuition	4,932	+85	75.4	1,572	+147	32.4	13,800	+117	60.5	1,500	+41	12.0
Model	2,906	-334	8.7	1,598	+138	1.5	26,964	+73	6.5	4,956	+107	2.1
Tax Filing Status	22,404	-73	71.0	10,888	+122	3.2	145,292	+59	69.0	18,550	+36	9.4

[1/] The frequencies are unduplicated totals within, but not between groups. For example, an applicant who made three corrections to adjusted gross income, twice pre-selection and once post-selection, is counted once in the pre-selection frequency and once in the post-selection frequency.

[2/] The SFI changes are for all corrections, solicited by comments and unsolicited, and are calculated from the time of correction to the most recent transaction.

2.10

All applicants tended to change AGI and TP to their disadvantage both pre- and post-selection/eligibility. Validation applicants made large corrections to their advantage to nearly all fields prior to selection; however, changes to AGI and TP, 90 percent of which came in response to edits, were to the applicants' disadvantage. The data strongly suggest that the numerous and restrictive edits which address AGI and TP are working effectively.

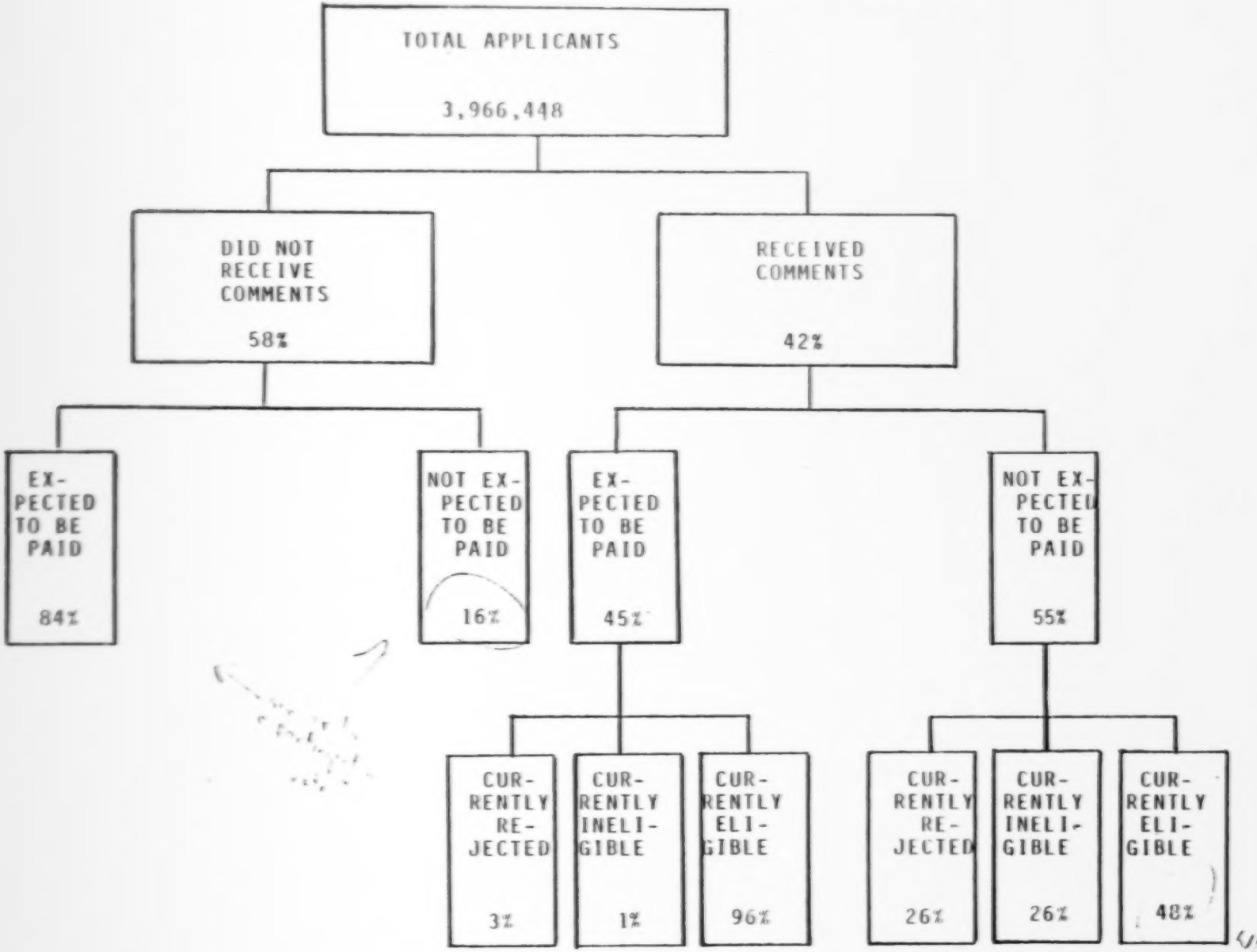
2.1.3 Current Eligibility and Payment Status of Applicants Receiving Comments

There is a concern that the processing system edits form needless roadblocks for many accurate and honest applicants. It is possible that some edits are too restrictive or too rigid to make allowances for special circumstances that do not meet the norm. Many applicants, discouraged or confused about the demands of the processing system, may be dropping out and never receiving a grant.

Exhibit 2.2 compares the current payment status of applicants who received comments addressing critical fields with those who did not. As a point of reference, applicants who are "expected to be paid" were on the recipient file as of May, 1980.

Exhibit 2.2 shows that the majority of applicants who received comments were not expected to be paid, while most applicants who did not receive a comment to a critical field were on the recipient file as of May, 1980. Given the May, 1980 deadline for making application corrections (with a much later deadline for validation applicants' corrections), it is possible that a certain number of those who are either currently rejected or ineligible and not expected to be paid will re-enter the processing system and receive a grant. Also, some who are currently eligible for a grant but are not on the recipient file may eventually receive a grant, owing to the late May and early June deadlines for applicants to submit their SERs for payment. However, it is equally plausible that many applicants in this comparatively large group--those receiving comments and not expected to be paid--have dropped out and will not receive a grant.

EXHIBIT 2.2: CURRENT PAYMENT AND ELIGIBILITY STATUS: APPLICANTS RECEIVING COMMENTS VS. APPLICANTS NOT RECEIVING COMMENTS



2.12

The following table presents the current payment status of applicants who received comments by the field that the comment addressed.

TABLE 2.4: CURRENT PAYMENT AND ELIGIBILITY STATUS BY CRITICAL FIELD OF COMMENT RECEIVED

CRITICAL FIELD:	NUMBER WHO RECEIVED COMMENTS TO FIELD	% EXPECTED TO BE PAID	NOT EXPECTED TO BE PAID		
			% REJECTED	% INELIGIBLE	% ELIGIBLE
AGI	1,009,936	44.1%	19.1%	11.0%	25.8%
TP	887,826	42.6	21.2	10.7	25.4
NTI	623,604	42.3	20.8	10.2	26.7
PORTIONS	709,186	41.2	21.6	10.9	26.3
VEB	143,998	39.0	12.9	20.7	27.4
HS	179,116	42.8	14.9	12.9	29.4
PHE	388,122	41.1	14.0	16.2	28.7
NA	191,796	48.0	11.3	20.9	19.7
UE	214,828	43.4	18.8	13.5	24.2
UT	68,908	39.9	22.0	13.8	24.3
AR	73,022	48.7	6.9	27.5	16.9
MS	113,316	39.7	21.3	13.1	25.8
MODEL	21,384	37.9	28.6	8.5	25.0
TFS	210,928	39.9	26.0	6.9	27.2

As Table 2.4 indicates, there is remarkably little variation in the current status of applicants by critical field comment received. Regardless of the field addressed by the comment, fewer than 50 percent of all applicants were expected to be paid and about 25 percent were eligible but had not received a grant. However, the current status of applicants receiving comments to NA, AR, TFS, and Model differs somewhat. Applicants receiving comments which addressed NA and AR were most likely to appear on the recipient file. The majority of those not expected to be paid who received comments to these two fields were currently ineligible; relatively few were currently rejected. This suggests that the edits and comments associated with these two fields are

not presenting any substantial barriers or discouraging applicants from ever receiving a grant. On the other hand, applicants receiving comments which address Model and TFS appear to be having trouble negotiating the system. Comparatively few of these applicants are expected to be paid, while over 25 percent are currently rejected.

2.1.4 Summary and Conclusions

The following summarize the key findings regarding the overall impact and performance of the processing system edits:

- About 42 percent of all applicants received at least one comment addressing a critical application item. Most of the comments generated addressed the following four key income fields: AGI, TP, Portions and NTI.
- One fourth of all applicants corrected a critical field. The majority of corrections were in response to processing system edits. Nearly one half of the solicited corrections did not cause a change in the applicants' SEI. Most of the corrections that changed the SEI were to the applicants' disadvantage.
- The processing system edits have a greater impact on applicants selected for validation than on applicants not selected. Validation applicants received comments at a greater rate than nonvalidation applicants. Plus, over 50 percent of validation applicants made pre-selection changes in response to the edits, whereas only 17 percent of nonvalidation applicants responded to the edits prior to attaining eligibility. It also appears that the edits have a greater impact than institution validation on the corrections behavior of validation applicants: only one-fourth made unsolicited post-selection corrections while over one-half corrected in response to the edits prior to selection.
- The following four key income fields were corrected most often: AGI, TP, Portions, and NTI. Least often corrected were: MS, VEB, Model, AR and UT. Over 85 percent of the corrections to AGI and TP were in response to a comment. This, together with the fact that corrections to AGI and TP tended to result in relatively large SEI changes to the applicants' disadvantage, indicate that the edits to these two fields are having a substantial impact on applicant corrections behavior.
- As of May, 1980, more than half of all applicants who had received comments were not expected to be paid, whereas the vast majority of those not receiving comments were expected to be paid. Comments relating to TFS and Model seemed to be causing applicants particular difficulty.

2.2: Rejection Edits and Comments

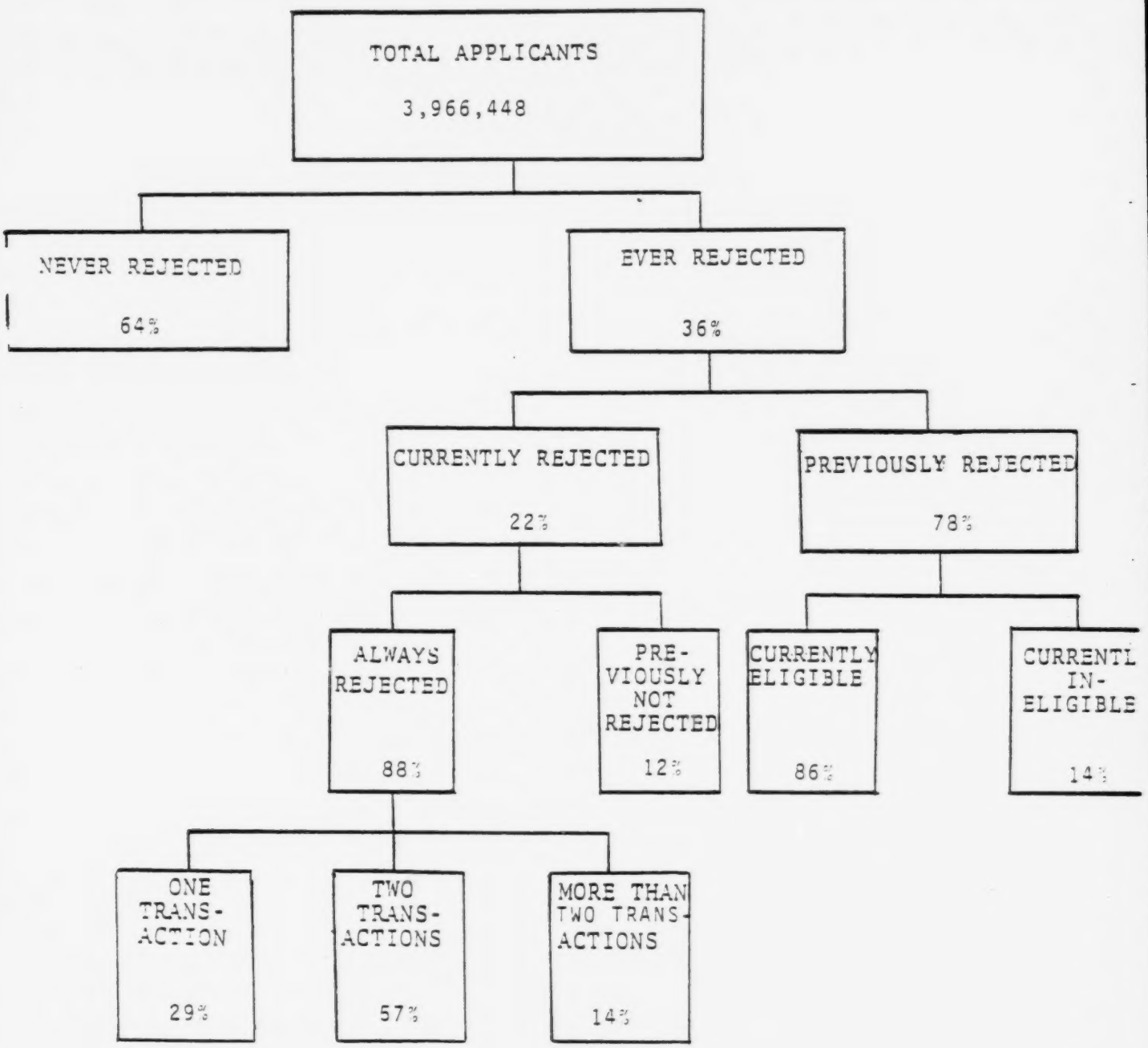
The most restrictive of the application processing edits are the rejection edits. When an application triggers a rejection edit, a comment is printed on the SER. In order to re-enter the processing system and receive an eligibility index, the rejected applicant must respond to the comment by either confirming that the item in question is correct or by correcting that item. There are 43 rejection edits, or rejection reasons, for applicants filing a regular application. Many rejection edits have more than one corresponding comment. The texts of the series of rejection comments that correspond to the same rejection reason are tailored to the applicant's marital and dependency status. For example, comments 10 and 12 are generated for the same reason, but to different populations: 10 to dependents with married parents and 12 to single independent applicants. This section analyzes the impact and the adequacy of the rejection edits and comments. Subsection 2.2.1 examines the income and household characteristics most affected by the rejection edits and comments, while Subsection 2.2.2 assesses the adequacy of the individual rejection edits and comments.

2.2.1 Characteristics of Applicants Most Frequently Rejected

As mentioned in Section 2.1, there is concern that many qualified and accurate filers are being impeded from receiving a grant because (1) the demands of the edits--the rejection edits in particular--are too severe and (2) the corrections process itself--that is, instruction given by the comments--is confusing. Exhibit 2.3, which follows, summarizes the impact of the rejection edits on the entire applicant population and shows the current eligibility status of applicants who had been rejected at least once.

The exhibit shows that over one-third of all applicants have, at one time, received a rejection edit. About 78 percent of these rejected applicants responded properly to the rejection edit or edits and, therefore, successfully re-entered the system and received an eligibility determination. The possibility exists that many of the 8 percent who are currently rejected will re-enter the system, having not yet responded to

EXHIBIT 2.3: SUMMARY OF CURRENT STATUS OF APPLICANTS EVER REJECTED



2.16

the rejection edit. The 208,278 individuals, or 5 percent of all applicants, who have remained rejected for two or more transactions are having the most difficulty in negotiating the system.

There is concern that the low income, needy applicants make up the majority of those applicants having the most difficulty negotiating the system and that the rejection edits may be categorically excluding this income group from the processing system. In order to address this concern, data on the status from previous to current transactions of applicants reporting various income and household data on their 1979-1980 applications was analyzed. The following table shows the income level of applicants most likely to become rejected and remain rejected.

TABLE 2.5: REJECTION RATES BY INCOME1/

INCOME LEVEL:	EVER REJECTED			NEVER REJECTED	Total Number of Applicants in Income Range
	Currently Rejected		Previously Rejected, Currently Eligible or Ineligible		
	Rejected One Transaction	Rejected Two or More Transactions			
Less than $0	3.1%	10.8%	81.6%	4.4%	9,936
0	11.0	28.9	60.0	.01	176,086
1 - 999	2.7	6.0	27.7	63.6	131,774
1,000 - 1,999	2.3	3.4	25.4	68.9	157,980
2,000 - 3,999	2.1	2.9	23.5	71.4	435,154
4,000 - 6,999	2.3	3.2	27.7	67.6	565,312
7,000 - 9,999	2.3	3.6	31.9	62.2	437,232
10,000 - 12,499	2.3	3.2	33.5	61.0	304,326
12,500 - 14,999	1.9	2.9	32.0	63.2	266,790
15,000 - 17,499	1.7	2.5	28.5	67.2	250,186
17,500 - 19,999	1.6	2.0	26.0	70.3	234,528
20,000 - 24,999	1.3	2.0	22.6	74.1	402,224
25,000 +	1.7	2.5	20.9	74.9	531,240

1/Income equals the sum of adjusted gross income and non-taxable income.

As Table 2.5 indicates, applicants reporting negative or zero incomes, who constituted about 5 percent of all applicants, became rejected and remained rejected at the highest rate. Nearly all zero and negative income applicants were rejected under rejection edit F (see Section 2.2.2 for a detailed examination of this edit), which screens all individuals reporting incomes less thant $51. However, once rejected, applicants with negative incomes had considerably more success than zero income applicants in re-entering the system. About 17 percent of the negative income applicants who ever received a rejectiom edit are currently rejected, whereas 68 percent of the zero income group who ever received a rejection edit are currently rejected. The rejection rates of the income groups between $1000 and $17,500 show little variation. Applicants with incomes from $1 to $999 become rejected and remain rejected slightly more frequently than applicants with incomes greater than $17,500.

The frequency of rejection of independents and dependents was also compared by income level in Table 2.6. The table shows that independent applicants were rejected and remained rejected at a slightly higher rate than dependents. However, the rejection rates of the two groups varied considerably by income level. Nearly all zero income applicants, independent and dependent, were ever rejected. Independents in this income group, however, had much more success in re-entering the processing system following rejection: as of May, 1980, only 24 percent of all zero income independents had been rejected for two or more transactions, while over 40 percent of their dependent counterparts had remained rejected for this length of time. In addition, independents with incomes from $1 to $15,000 were rejected considerably less often than dependents in this income range. For example, nearly half of all dependents reporting an income between $1 and $999 were rejected, whereas only one-third of their independent counterparts ever received a rejection edit. In sum, the rejection edits are having a greater degree of impact on dependents with very low incomes than on the more numerous independents in this income group.

TABLE 2.6: REJECTION RATES BY DEPENDENCY STATUS AND INCOME

Income Level 1/	Independent			Dependent		
	# of Applicants	% Ever Rejected	% Remained Rejected Two or More Transactions	# of Applicants	% Ever Rejected	% Remained Rejected Two or More Transactions
Less than $0	1,070	96.1%	20.9%	8,866	95.5%	9.6%
0	129,472	99.9	24.3	46,614	99.9	41.8
1 - 999	102,046	32.7	4.1	29,728	49.3	12.7
1,000 - 1,999	115,072	28.3	2.7	42,908	38.6	5.4
2,000 - 3,999	274,412	25.2	2.6	160,742	34.4	3.5
4,000 - 6,999	278,412	28.1	3.4	287,064	36.5	3.1
7,000 - 9,999	144,814	31.7	4.5	292,418	40.8	3.2
10,000 - 12,499	66,760	30.8	4.9	237,566	41.3	2.8
12,500 - 14,999	40,232	30.5	4.6	226,558	37.9	2.6
15,000 - 17,499	28,140	30.1	4.6	222,046	33.1	2.2
17,500 - 19,999	18,288	27.4	4.8	216,240	29.9	1.8
20,000 - 24,999	19,780	24.7	4.3	382,464	26.0	1.8
25,000 +	13,608	34.3	7.6	517,632	24.9	2.4
Population Average	1,242,646	36.7	6.3	2,723,802	35.5	4.7

1/Income equals the sum of adjusted gross income and nontaxable income.

The following table lists the application charateristics, in addition to a very low income, which are most likely to trigger a rejection edit:

TABLE 2.7: CHARACTERISTICS OF APPLICATIONS MOST FREQUENTLY REJECTED

APPLICATION CHARACTERISTICS	NUMBER OF APPLICANTS WITH CHARACTERISTIC	% EVER REJECTED	% REMAINED REJECTED TWO OR MORE TRANSACTIONS
MS - DECEASED	1,780	100.0%	71.6%
UE - $6,000+	9,084	100.0	18.8
AGI - BLANK	82,586	97.9	73.3
HS - BLANK	36,166	95.2	67.0
PHE - BLANK	31,688	94.5	67.4
VEB - $7,000+	1,616	75.2	34.7
VEB - $1,001-2,000	80,102	63.7	7.6
TP - BLANK	145,610	61.3	36.0
NTI - $12,500+	28,486	57.8	8.7
UT - BLANK	541,048	51.5	13.1
PHE - 5+	12,698	50.6	13.5
MS - WIDOWED	273,914	50.5	5.4

The processing system, as Table 2.7 shows, automatically rejects applicants who file as dependents and report that both parents are deceased at the time of the original application. As of May, 1980, none of the applicants in this comparatively small group had re-entered the system and obtained an eligibility determination; 72 percent, as the table indicates, had been rejected for two or more transactions. The 1979-80 Basic Grant application does not inform the applicant that he or she is an independent if both parents are deceased at the time of the first filing, even though one or both parents may have provided assistance, claimed the applicant as a tax exemption, or provided a home for the applicant during 1978. The 1980-81 applicant resolved this problem: its instructions clearly state that the applicant is automatically an independent if both parents are deceased.

Table 2.7 also indicates that all applicants reporting unusual expenses greater than $6,000--once again, a comparatively small number--are automatically rejected. Section 2.2.2 provides a detailed examination of the rejection edit which screens applicants reporting a large amount of unusual expenses.

Applicants leaving key income and household items blank have a very high likelihood of becoming rejected and remaining rejected. Also, as Table 2.7 indicates, the processing system frequently rejects applicants reporting veteran's benefits between $1,000 and $2,000. In fact, applicants reporting any amount of veteran's educational benefits were rejected at a notably higher rate than applicants not reporting benefits: 35 percent of applicants with no benefits were ever rejected, whereas the processing system rejected 50 percent of the applicants reporting veteran's benefits.

2.2.2 Adequacy of the Individual Rejection Edits and Comments

This subsection assesses the adequacy of the individual rejection edits and comments by examining applicant corrections and the ability of applicants to re-enter the processing system following the receipt of a rejection edit. More specifically, the following criteria were used to analyze applicant response to each edit and comment:

- frequency of corrections
- frequency of verifications
- magnitude and direction of corrections measured by effective SEI change and potential payment change
- ability of applicants to re-enter the system following corrections
- current payment status of applicants who received rejection edits

Ideally, a rejection edit should correctly identify those with missing or inaccurate data. The directions given by the rejection comment should be clear and easily understood to insure a complete, accurate and prompt response. With this in mind, the following assumptions were made in labeling a rejection edit "successful":

- A high percentage of applicants will change suspect data items in response to a successful edit. Likewise, few will verify that the data in question is correct.
- In order to be worthwhile, an edit should solicit corrections that have more than a negligible impact on the applicant's SEI and potential grant size.
- A high percentage of applicants should be able to obtain an eligibility determination on the transaction immediately following receipt of a successful rejection edit. Thus, the majority of applicants who received the successful edit should currently be eligible or ineligible.

Applicants can be rejected for more than one reason on a single transaction. The system, however, notes only one of the reasons, or codes, even though all applicable rejection comments are generated on the SER. The codes for each rejection edit are given the priority order alpha A to Z then numeric 1 to 17. For example, if an applicant is rejected for reasons B and 7, the system notes only code B but the student receives comments for both reasons. In order to gain a more accurate understanding of the adequacy of the rejection reasons, the data for the analysis in this section is based on the actual reasons for rejection--that is, the rejection comments that were generated--and not on the priority rejection code appearing on the applicant file. The rejection code, as used in this analysis, provides a convenient way to group comments by rejection reason.

Some caveats concerning the interpretation of the findings in this section must be stated:

1) Applicants who received edit comments and who make changes to the fields to which the comments apply may be submitting these corrections for reasons other than those solicited by the edit comments they received. To the extent that corrections which appear to be made in response to the edit comments are made for other reasons, the attribution of eligibility index changes to the edits may be overestimated.

2) Applicants who are rejected do not have an eligibility index. In order to assess the impact of the edits, it was necessary to compute an eligibility index for applicants on the transaction of rejection and compare the difference between that index and the index received after they responded to the rejection comment.

The computation of the index is confounded by the fact that one of the reasons for rejection is missing data. Ideally, the way to obtain reliable estimates for missing data is to conduct a nonrespondent survey. In this case, it would entail selecting a representative sample of rejected applicants with missing data, and contacting them to find out true values. Unfortunately, this study is confined solely to statistical analysis of existing data, and it is not within the scope of this contract to conduct such a survey. Two alternative methodologies that are limited to using existing data were considered for this computation. One was to substitute average values from applicants who had been rejected, but provided data, for those who had missing data. The second was to assume values based on other data reported, where possible, and where not possible, to assume the missing value to be zero.

Neither of these alternative procedures is ideal because each introduces errors. However, given the study goals, a procedure was selected that was judged to have the least serious limitations.

The limitations of the method of substituting average values are:

a) The fact that the applicants (whose average values could be used for substitution) provided data made them different from those who omitted data. Therefore, from a statistical design perspective, these applicants are not equivalent and such a substitution would introduce biases of an unknown magnitude and direction.

b) From an operational perspective, the number of categories of applicants who would have to be defined would be extremely large, greatly increasing the time necessary to process the computations.

The limitation of the zero/logically derived substitution approach is that it introduces a bias of an unknown magnitude, but a known direction. That is, we know that we may be using a zero instead of a value. Therefore, this procedure introduces a bias which probably understates actual values to some unknown degree.

The approach that was selected was the zero/logically derived substitution approach because the unknown component of the error is limited to magnitude and not direction. The specific assumptions used for each field can be found in Appendix B.

As mentioned previously, there are 43 rejection edits for nonsupplemental applicants. For this analysis, however, only those edits which address critical fields were examined. Table 2.8 provides a list

of the number of rejected applicants by rejection reason. The following two groups summarize the most frequent and least frequent rejection reasons:

Most Frequent		Least Frequent	
G -	Reported tax greater than computed tax by $300 and 50 percent of computed tax	C -	HS greater than 19
		A -	Reported tax greater than 50 percent AGI
D -	Portions greater than 120 percent or 170 Percent AGI	K -	Debts greater than assets - farm
F -	Zero, negative or less than $51 income	R -	Applicant's resources decreased by $300

In addition to the above, a relatively large number of applicants received one of the following four rejection edits, given for leaving key application fields blank: 7, 8, 16 and 17. In particular, many applicants reported at least one earned income portion, but left AGI blank. Also, a surprisingly large number of applicants were rejected for reporting unusual expenses of greater than $5,000.

Table 2.9, which follows Table 2.8, shows the percentage of all rejected applicants correcting in response to a rejection comment for each of the rejection reasons. The most successful edits in terms of the frequency of solicited corrections were reasons S and T, generated when the amount of social security benefits reported on the Basic Grant application does not match the amount recorded on the Social Security Administration's computer file. Eighty percent of all applicants rejected for reasons S and T corrected NTI on the first subsequent transaction.

As would be expected, a high percentage of applicants corrected in response to the four edits associated with missing data. In particular, reason 7, which screens applications where both MS and HS are blank, solicited a high rate of corrections. Remarkably few changed an appropriate field in response to reasons F - zero, negative or less than $51 income - and E - UE greater than income or $5,000. This strongly

TABLE 2.8: DISTRIBUTION OF REJECTED APPLICANTS BY REJECTION REASON AND CORRESPONDING COMMENT

Code/Rejection Reason	Comments	Number Who Received[1/]
A - Reported tax greater than 50% AGI	259	1,790
B - Combination verify	- -	114,184
	10	61,122
	12	17,090
	13	12,264
	128	22,984
	247	16
	297	4
	298	370
	299	334
C - HS greater than 18	18	506
D - Portions greater than 120% or 170% AGI	- -	190,280
	27	84,828
	41	27,544
	42	44,224
	234	33,684
E - Unusual expenses greater than income or greater than $5,000	30	109,328
F - Zero, negative, or less than $51 income	- -	173,926
	58	21,170
	98	106,390
	99	18,120
	238	27,370
	300	294
	301	96
	302	20
	303	466
G - Reported tax greater than computed tax by $300 and 50% of computed tax	47	210,726
H - Debts greater than assets - home	179	47,442
I - Debts greater than assets - investments	180	17,130

2.25

TABLE 2.8: (Continued)

Code/Rejection Reason	Comments	Number Who Received[1/]
J - Debts greater than assets - business	181	22,880
K - Debts greater than assets - farm	182	6,244
L - Unreimbursed tuition greater than 35% of income	212	25,778
R - Applicant's resources decreased by $300	256	7,044
S - SSA match - blank or zeroes	- -	39,096
	138	35,570
	287	460
	288	3,066
T - SSA match - SS benefits reported	- -	50,398
	139	48,336
	289	238
	290	1,824
U - VA match - blank or zeroes	283	14,422
V - VA reported and less than $131	284	34,182
7 - MS, HS Blank	- -	24,110
	70	14,314
	140	9,796
3 - NTI, AGI and Portions all blank	- -	27,890
	28	12,580
	43	4,970
	44	1,528
	236	8,400
	307	198
	309	196
	310	16
	311	2

TABLE 2.8: (Continued)

Code/Rejection Reason	Comments	Number Who Received[1/]
9 - Divorced, widowed or unmarried comments	- -	50,766
	24	0
	33	31,130
	35	18,240
	106	1,280
	148	0
	270	0
	308	116
	313	0
16 - AGI blank and portion reported	- -	146,658
	268	0
	271	78,502
	273	36,376
	275	8,728
	277	21,646
	306	6
	315	402
	316	504
17 - TP blank and AGI greater than zero	- -	37,164
	251	4
	269	48
	279	18,726
	280	8,364
	281	2,968
	282	6,132
	312	588
	314	334

[1/]The frequencies are duplicated, given that many applicants received more than one rejection reason.

TABLE 2.9: PERCENTAGE OF APPLICANTS CORRECTING RELEVANT CRITICAL FIELDS ON THE [illegible] SUBSEQUENT TO RECEIVING A REJECTION EDIT

REJECT REASON CODE	CRITICAL FIELD: Adjusted Gross Income	Taxes Paid	Positions	Nontaxable Income	Veteran's Educational Benefits	Net Assets	Marital Status	Household Size	Unusual Expense	Unreimbursed Tuition	Applicant's Resources	Total Number Who Received Edit
A	44%1/	49%	--	--	--	--	--	--	--	--	--	1,790
B	63	38	24	16	--	--	--	--	--	--	--	114,184
C	--	--	--	--	--	--	--	59	--	--	--	506
D	37	14	21	--	--	--	--	--	--	--	--	190,279
E	--	--	--	--	--	--	--	--	22	--	--	109,328
F	14	5	10	13	--	--	--	--	--	--	--	173,926
G	32	52	--	--	--	--	--	--	--	--	--	210,726
H	--	--	--	--	--	74	--	--	--	--	--	47,442
I	--	--	--	--	--	68	--	--	--	--	--	17,130
J	--	--	--	--	--	30	--	--	--	--	--	22,880
K	--	--	--	--	--	51	--	--	--	--	--	6,244
L	--	--	--	--	--	--	--	--	--	37	--	25,778
R	--	--	--	--	--	--	--	--	--	--	27	3,066
S	--	--	--	80	--	--	--	--	--	--	--	39,096
T	--	--	--	80	--	--	--	--	--	--	--	50,398
U	--	--	--	--	43	--	--	--	--	--	--	14,422
V	--	--	--	--	35	--	--	--	--	--	--	34,182
7	--	--	--	--	--	--	80	72	--	--	--	24,110
8	66	54	56	61	--	--	--	--	--	--	--	27,890
9	43	35	17	16	--	--	24	--	17	6	--	50,650
16	79	47	32	--	--	--	--	--	--	--	--	146,658
17	45	77	40	38	--	--	--	--	--	--	--	37,164

1/Percentages are based on the total number of applicants who received the reject reason. For example, 44 percent of the 1790 applicants who received reject reason A corrected adjusted gross income on the first subsequent transaction. Percentages within reject reasons are based on duplicated frequencies. An applicant rejected for reason A, for example, might correct both adjusted gross income and taxes paid, and thus would be counted in both percentages.

suggests that these two edits, both of which were given to a comparatively large number of applicants, are not successfully identifying applicants with inaccurate data. Reason G - reported tax greater than computed tax by $300 and 50 percent of computed tax - is worth examining closely since the highest percentage of rejected applicants met this reason. It appears to be effectively identifying applicants who misreported TP, with 52 percent correcting this application item; to a lesser extent, reason 6 identified applicants who apparently misreported AGI, with 32 percent correcting this field.

Besides changing a critical field, a rejected applicant has the option of confirming that the data in question is correct. Table 2.10 shows the percentage of applicants, by rejection reason, who verified critical fields on the transaction just subsequent to receiving the rejection reason.

TABLE 2.10: VERIFICATION BY REJECTION REASON

REJECT REASON CODE	NUMBER WHO RECEIVED EDIT	% WHO VERIFIED	REJECT REASON CODE	NUMBER WHO RECEIVED EDIT	% WHO VERIFIED
A	1,790	6.14%	L	25,778	28.32%
B	114,184	5.89	R	3,066	39.52
C	506	13.19	S	39,096	0.02
D	190,279	22.94	T	50,398	0
E	109,328	50.32	U	14,422	0
F	173,926	46.79	V	34,182	0
G	210,726	14.67	7	24,110	0
H	47,442	9.85	8	27,890	0
I	17,130	15.59	9	50,650	0
J	22,880	45.80	16	146,658	0
K	6,244	33.28	17	37,164	0.04

An inverse relation exists between the percent who corrected and the percent who verified in response to a rejection edit. A very high percentage verified questioned data while relatively few corrected in

response to reasons E and F. A high proportion - 74 percent - corrected in response to reason H - home debts greater than assets - while comparatively few verified. Not surprisingly, almost none verified in response to the four edits which request that missing information be provided.

In order to be worthwhile, an edit has to solicit corrections that have more than a negligible impact on the applicant's potential award amount. Table 2.11, which follows, shows the magnitude and direction of applicants' corrections to critical fields in response to the rejection edits. Before examining this table, refer again to the caveats regarding the interpretation of the data outlined earlier in this section. The payment difference in Table 2.11 is based on the potential expected disbursement from the time the edit was received to the first subsequent transaction. As a point of reference, the expected disbursement is the amount of award an applicant is due to receive, taking into account enrollment status, cost of education and the SEI. (In the case of the rejected applicant, the expected disbursement is based on the specially computed SEI.)

In general, as Table 2.11 indicates, about half of the corrections solicited by the rejection edits did not have an impact on the applicant's SEI. The 50 percent who changed their SEI as a result of a correction tended to reduce their expected award between 50 and 300 dollars. However, the corrections in response to the following six edits differed notably from the above generalizations: F, H, S, T, 7, and 9. Over 70 percent of the corrections in response to reasons F - zero, negative or less than $51 income - and H - home debts greater than assets - had no impact on the applicants' SEI. In contrast, fewer than 10 percent of the corrections to S and T - no social security match - were inconsequential in terms of the SEI. Over 80 percent responded to reasons S and T with corrections to their disadvantage. The data strongly suggest that the two social security match edits are consistently identifying misreporting, whereas reason F, given to applicants reporting very low income, does not appear to be identifying inaccurate information.

TABLE 2.11: THE MAGNITUDE AND DIRECTION OF CORRECTIONS IN RESPONSE TO REJECTION EDITS IN TERMS OF EFFECTIVE SEI CHANGE AND AVERAGE POTENTIAL PAYMENT DIFFERENCE

Reject Reason Code	Percent with Positive SEI Change	Average SEI Increase	Percent With Negative SEI Change	Average SEI Decrease	Percent with No SEI Change	Overall Average SEI Change	Overall Potential Payment Change	Total Number of Corrections
A	26%	+565	7%	-847	67%	+141	-$42	1,680
B	70	+817	3	-510	27	+555	-$392	170,010
C	57	+884	5	-593	38	+472	-$362	302
D	27	+666	6	-498	67	+152	-$82	153,236
E	48	+803	2	-830	50	+367	-$243	28,222
F	27	+852	1	-425	72	+228	-$300	76,522
G	53	+365	9	-434	38	+156	-$109	196,414
H	22	+434	7	-546	71	+58	-$40	37,044
I	29	+441	7	-574	64	+87	-$75	12,368
J	25	+391	7	-486	68	+63	-$55	8,694
K	24	+575	7	-599	69	+99	-$87	3,460
L	44	+789	1	-336	55	+341	-$239	10,356
R	74	+356	1	-680	25	+251	-182	837
S	86	+323	5	-341	9	+261	-$174	31,444
T	82	+263	9	-224	8	+195	-$130	40,608
U	38	+494	5	-910	58	+144	-$73	6,184
V	31	+264	25	-157	44	+43	-$23	12,986
7	9	+661	29	-831	62	-185	+$128	29,016
8	52	+922	1	-676	47	+479	-$353	65,820
9	11	+457	48	-1221	41	-536	+$377	36,844
16	38	+340	19	-366	43	+60	-$44	316,122
17	41	+291	28	-451	31	-10	-$7	74,132
TOTAL OTHER NON-REJECTION EDITS	30	+455	16	-447	54	+62	-$17	272,796

1/SEI changes and payment differences are based on the total number of corrections to related critical fields. Note that they are based on the number of corrections, not on the number of applicants making corrections.

All the edits, except reason 7 - HS and MS blank - and 9 - number of portions and MS inconsistent - solicited corrections resulting in lower award potential. It should be noted that it is the function of both edits to cause the award level to increase. Applicants who meet reason 7 will automatically increase their eligibility by providing the missing household size figure. Reason 9 rejects applicants whose marital status is inconsistent with the number of portions reported; the data indicate that applicants are unnecessarily reporting the income of a second person.

Corrections to the following rejections resulted in relatively small potential award changes:

H - home debts greater than assets
V - VEB reported and less than $131
17 - TP blank and AGI greater than zero

The ability of applicants to re-enter the system and receive an eligibility determination following a correction in response to a rejection edit is illustrated in Table 2.12. As this table indicates, applicants responding to reason C - HS greater than 19 - were most successful in re-entering the system, with only 9 percent remaining rejected on the first subsequent transaction. Applicants responding to the social security and veteran's educational benefits edits (S, T, U and V) and to reason 7 - HS and MS blank - were also relatively successful in re-entering the system. In fact, applicants responding to edits that required corrections to only one or two fields tended to remain rejected less frequently than applicants who responded to edits requiring corrections to three or more fields. A comparatively high proportion of applicants who corrected in response to reason 9, an edit which demands corrections to seven fields, remained rejected. This suggests that either the demands of this edit are too severe or that the instructions given by the corresponding comments are unclear and confusing.

Table 2.13 shows the current eligibility and payment status of applicants who have ever been rejected by the rejection reason. This table indicates the degree to which each edit may have impeded applicants

TABLE 2.12: PERCENTAGE OF APPLICANTS WHO REMAIN REJECTED ON THE TRANSACTION JUST SUBSEQUENT TO MAKING A CORRECTION TO A FIELD RELEVANT TO A REJECT REASON CODE

	CRITICAL FIELD										
REJECT REASON CODE	Adjusted Gross Income	Taxes Paid	Portions	Nontaxable Income	Veteran's Educational Benefits	Net Assets	Marital Status	Household Size	Unusual Expense	Unreimbursed Tuition	Applicant's Resources
A	32%1/	29%	--	--	--	--	--	--	--	--	--
B	23	32	31	35	--	--	--	--	--	--	--
C	--	--	--	--	--	--	--	9	--	--	--
D	22	33	27	--	--	--	--	--	--	--	--
E	--	--	--	--	--	--	--	--	32	--	--
F	23	32	36	18	--	--	--	--	--	--	--
G	28	21	--	--	--	--	--	--	--	--	--
H	--	--	--	--	--	19	--	--	--	--	--
I	--	--	--	--	--	20	--	--	--	--	--
J	--	--	--	--	--	34	--	--	--	--	--
K	--	--	--	--	--	26	--	--	--	--	--
L	--	--	--	--	--	--	--	--	--	27	--
R	--	--	--	--	--	--	--	--	--	--	20
S	--	--	--	15	--	--	--	--	--	--	--
T	--	--	--	14	--	--	--	--	--	--	--
U	--	--	--	--	15	--	--	--	--	--	--
V	--	--	--	--	21	--	--	--	--	--	--
7	--	--	--	--	--	--	16	19	--	--	--
8	35	38	41	38	--	--	--	--	--	--	--
9	43	46	42	41	--	--	34	--	49	56	--
16	25	30	31	--	--	--	--	--	--	--	--
17	25	22	24	24	--	--	--	--	--	--	--

1/Percentages are based on the number of applicants who received the reject reason and corrected the corresponding field. For example, 32 percent of the applicants who received reason A and corrected adjusted gross income remained rejected on the first subsequent transaction.

TABLE 2.13: CURRENT ELIGIBILITY AND PAYMENT STATUS OF EVER REJECTED APPLICANTS BY REJECT REASON

Reject Reason Code	Not Expected To Be Paid				Expected To Be Paid				Total Number who Received Edit
	% Not Expected To Be Paid	% Currently Rejected	% Currently Ineligible	% Currently Eligible	% Expected To Be Paid	% Currently Rejected	% Currently Ineligible	% Currently Eligible	
A	60% [1/]	28%	8%	24%	40%	2.0%	0	38%	1,790
B	61	22	14	25	39	0.3	0.2	38	114,184
C	58	28	15	15	42	0.4	0	42	506
D	53	18	8	27	47	0.4	0.1	46	190,280
E	55	17	11	27	45	0.4	0.1	45	109,328
F	61	23	2	36	39	0.3	0.1	39	173,926
G	55	20	13	22	45	0.3	0.1	44	210,726
H	52	14	15	23	48	0.4	0.1	47	47,442
I	59	15	27	17	41	0.2	0.1	41	17,130
J	51	11	20	20	49	0.1	0.1	49	22,880
K	45	9	16	20	55	0.1	0	55	6,244
L	55	18	9	28	45	0.4	0.1	45	25,778
R	55	11	17	27	45	0.8	0.5	43	7,044
S	45	12	7	26	55	0.1	0.1	55	39,096
T	39	9	6	24	61	0.1	0.1	61	50,398
U	71	32	9	30	29	0.1	0.1	29	14,422
V	44	5	20	19	56	0.2	0.1	56	34,182
7	70	28	43	29	30	0.6	0.1	30	24,110
8	68	34	12	22	32	0.3	0.1	32	27,890
9	87	48	7	32	13	0.6	0.2	12	50,766
16	59	23	16	20	41	0.3	0.1	41	146,658
17	64	26	18	20	36	0.1	0.1	36	37,164
Total Other Non-Rejection Edits	53	10	14	29	47	0.2	0.1	47	1,206,878

1/Percentages are based on the total number who received the edit. The percentage not expected to be paid plus the percentage expected to be paid equals 100 percent; the sum of the other 6 columns equals 100 percent.

from receiving a grant. In general, the data in Table 2.13 corroborates the findings from Table 2.12. Applicants rejected for reasons S, T, and V, all of which correspond to the social security and veteran's benefits' tape matches, were most likely to appear on the recipient file as of May, 1980. Fewer than forty-five percent of the applicants who received one of these three edits were not expected to be paid. In contrast, 53 percent of the applicants who received non-rejection edits were not expected to be paid. This indicates that rejection reasons S, T, and V created no significant roadblocks for rejected applicants.

On the other hand, reason 9 appears to be a substantial barrier: 87 percent of those who received this rejection edit were not expected to be paid. A high proportion of applicants meeting reason 7 were also not expected to be paid. However, the majority in this group were currently ineligible, rather than currently rejected as most were who had met reason 9.

The current status of applicants rejected for reason F - zero, negative or less than $51 income - is worth examining closely since a large percentage of applicants in this group are currently eligible and expected to be paid. This either indicates that many applicants have not been recorded on the file or that a large number "dropped out" and did not submit an eligible SER to their financial aid administrator.

2.2.3 Summary and Conclusions

The following summarizes the key findings regarding the impact and adequacy of the rejection edits:

- About 36 percent of all applicants were rejected by the processing system for providing missing or apparently inaccurate information. As of May, 1980, over three-fourths of the rejected applicants had responded properly to the rejection comments and were given an eligibility determination. The majority of applicants who were currently rejected had been rejected on two or more transactions. In general, the rejection edits form a greater roadblock than the other less restrictive edits, with a higher percentage of applicants currently not expected to be paid who received rejection edits than those who did not. A majority of the corrections in response to the rejection edits had no effect on the applicants' SEI. Those corrections that did have an impact tended to be to the applicants' disadvantage.

- Applicants reporting negative and zero incomes were rejected at a significantly higher rate than the rest of the population. Negative income applicants appeared to have little difficulty re-entering the system following rejection; however, a large percentage of zero income individuals, in particular the dependents, remained rejected for two or more transactions. Edit F, which rejects applicants reporting an income less than $51, fared poorly by all the criteria used in this analysis. Very few applicants corrected in response to reason F. Three-fourths of those who corrected did not change their award potential. Nearly half responded to reason F by confirming that the data in question were correct. The data strongly suggest that the processing system is creating an unnecessary roadblock for very low income applicants, given that it appears that reason F is not soliciting award reductions. Conversely, the data could mean that reason F is a strong deterrent for the misreporters. Clearly, further study is needed regarding the accuracy of very low income applicants. If further study shows that applicants reporting negative, zero or less than $51 income do not misreport more often than other applicants then consideration should be given to abolishing edit edit F. Otherwise, the following refinements could be made to the edit to make it less restrictive:

  -- Do not reject applicants with negative incomes, given that the 1974-75 and 1976-77 IRS studies showed that negative income applicants are among the most accurate reporters of income and taxes;

  -- Match low income with household size or number in post-high enrollment. For example, reject applicants with one in the household and income less than $500, two in household income less than $1000, etc.

- In addition to very low income, missing data, taxes inconsistent with income and portions inconsistent with AGI were major reasons for rejection. About 6 percent of all applicants were rejected for leaving a key household or income item blank. Most applicants in this group reported an earned income portion, but left AGI blank. In general, the applicants with missing application information had difficulty entering the system, with over 60 percent remaining rejected for two or more transactions. The data indicates the need for continued emphasis in the application instructions of the necessity of providing complete information.

- In addition to rejection reason F, reason E - UE greater than income or $5,000 - and reason 9 - number of portion inconsistent with marital status - performed poorly. Reason E, received by 3 percent of all applicants, solicited few corrections and had a high percentage of verifications, suggesting that many applicants do in fact report unusual expenses greater than

$5,000 even after being questioned by the system. With more higher income applicants than in prior years and because of the general impact of inflation (i.e., increased medical and dental expenses), it follows that many applicants have expenses greater than $5,000. It appears, then, that reason E should be refined. Making this edit less restrictive by raising the $5,000 level would be the most logical refinement. Nearly 90 percent of the 50,766 applicants who met reason 9 were not on the recipient file as of May, 1980. A comparatively high percentage remained rejected on the first transaction after receiving edit 9. The texts of the comments corresponding to this edit are long; all demand corrections to seven fields. The data suggest that the comments' instructions, because of their length and relative complexity, are confusing applicants and impeding many from re-entering the system. Conversely, it could mean that the applicant misreported and concluded that it was not worthwhile to re-enter the system.

- The edits associated with the Social Security and Veteran's benefit tape matches are the most effective of the rejection edits based on the criteria used in this analysis. A relatively high percentage corrected (particularly in response to the Social Security edits), almost none verified that the data in question were correct, and a very high percentage promptly re-entered the system following the receipt of the edit.

3

ADEQUACY OF THE PRE-ESTABLISHED CRITERIA

This Chapter assesses the efficacy of the the pre-established criteria as indicators of misreporting. The Chapter is divided into four major sections. The first section (3.1) provides background information on the PEC and describes the study design and the measures of effectiveness analyzed. The second section (3.2) attempts to determine whether selecting applicants for validation because they meet PEC, generally identifies applicants likely to misreport better than random selection. Section 3.3 examines the relative efficacy of the PEC subgroups; the final section (3.4) summarizes the findings and presents recommendations.

3.1 Background on the Pre-established Criteria

Several years ago, the Bureau of Student Financial Assistance developed a set pre-established criteria (PEC) to help identify students likely to misreport information on their BEOG application. Misreporting is here defined as the occurrence of post-selection (eligibility) corrections which result in SEI changes to the applicant's disadvantage. The criteria, based on previous analyses of misreporters, established categories of applicants with questionable information on their application and/or suspicious corrections behavior, and targeted a certain number of applicants in each category for validation.

The PEC have traditionally consisted of three groups: the A group is intended to identify applicants who have been rejected for reporting incomplete or inconsistent information, and who, through one or more

corrections, have established eligibility; the B group relates to applicants who were previously ineligible, and who, through one correction have significantly reduced their eligibility index and are now eligible; the C group identifies eligible applicants who significantly reduce their highest eligibility index with one, continuous, official series of transactions. Each PEC group is made up of several subgroups or criteria which more finely specify suspect applicant characteristics and behavior. In 1979-80, the number of PEC subgroups doubled from the number used in the pevious year. Eleven new subgroups were added to criterion A, three to criterion B, one subgroup to criteria C, and a new criterion group D (with 3 subgroups). The D group identifies applicants who report having four or more family members in post-high-school education and/or applicants whose reported financial data is estimated instead of based on actual tax information. The new A subgroups are concerned primarily with Veteran and Social Security Benefits and crossing B, C, and D criteria with A criteria; the new B and C criteria allow for crossing all possible combinations of B, C, and D criteria. Appendix D at the end of this report lists the PEC used in 1979-80 and explains their meaning.

As of May, 1980, the cut-off for data for this report, 166,348 applicants were selected for validation because they met the pre-established criteria; 46,014 or approximately 22 percent of the applicants validated were chosen randomly. Table 3.1 presents a summary of the number of validation applicants selected according to pre-established criteria and randomly.

TABLE 3.1: SUMMARY OF VALIDATION CEILINGS AND NUMBER OF APPLICANTS SELECTED

Criterion Group	Final Ceiling	# of Applicants Selected for Validation	% of Validation Population
A	127,000	137,618	45.9
B	25,000	21,140	7.1
C	17,500	13,274	4.4
D	70,000	81,266	27.2
Random	60,000	46,014	15.4
Total -	299,500	299,312	100

As can be seen in the above table, the majority of students chosen for validation were selected because they met one of the A criteria.

3.1.2 Study Design

The selection of validation applicants according to two major groups - those meeting PEC and those selected randomly, as well as the division of those meeting PEC into their appropriate subgroups, facilitates comparisons which can indicate the relative effectiveness of each group (and subgroup) in identifying misreporters. In these analyses, effectiveness is compared at several levels. First, measures of effectiveness for the random validation group are compared with the same measures of effectiveness for the total group of validation applicants meeting PEC to determine if the PEC criteria in general are superior indicators of misreporting. Next, measures of effectiveness for the PEC criteria groups (A, B, C, D) are compared by individual groupings and by measures of effectiveness for all PEC validation applicants to determine the relative efficacy of each group. Finally, the specific criterion subgroups within each PEC grouping are compared to each other.

3.1.3 Measures of Effectiveness

Previous studies (including Phase I of this study) have attempted to assess the efficacy of the PEC primarily through the use of two measures:

- the percentage of applicants submitting post-selection corrections; and
- the magnitude of the SEI changes resulting from these corrections.

The use of the measure "rate of applicants correcting post-selection" is based on the assumption that the validation process detects applicants who have reported inaccurate information on their application, and causes them to correct the inaccuracies. Corrections, therefore, are synonomous with misreporting. The higher the rate of applicants correcting, the greater the number of misreporters identified.[1/] The use of the measure "magnitude of SEI change resulting from corrections" is a necessary complement to the rate of corrections in judging effectiveness since the value of corrections which result in little or no SEI change is questionable at best. This measure is based in part on the assumption that the degree of SEI change reflects the magnitude of the discrepancy between previously incorrect and now correct information. Therefore, it follows that the greater the magnitude of SEI change, the more effective the criteria in identifying misreporters.[2/]

[1/] It is assumed that the new post-selection corrections represent accurate information. Although study findings support this assumption (i.e., post-selection corrections tend to reduce awards), it is an untested assumption because no data are available on the extent to which institutions carry out the validation procedures, and the degree to which data are correctly verified. A study of institutionally-verified income and taxes is currently being undertaken in cooperation with the Internal Revenue Service. This should provide some insight into the extent to which institutions have validated these data.

[2/] For these analyses, effectiveness of the PEC is determined by the magnitude of the SEI change, regardless of the direction of the change. It is assumed that since the correct award of Basic Grants is one of the program goals, it is appropriate for individuals both overreporting and underreporting to be corrected.

While the rate of applicants correcting post-selection and the magnitude of the resulting SEI change are the primary measures of effectiveness, we have also used several other measures to address effectiveness. These measures provide added depth to the analyses and allow us to focus on the impact of each PEC group in detail. These measures are as follows:

- the percentage of applicants correcting post-selection whose corrections do not result in any SEI change (in contrast to the gross measure above of applicants correcting post-selecting regardless of whether the correction elicits SEI change);
- the relationship between the reason a student was chosen for validation and the critical fields the applicant corrects;
- the relationship between the percentage of applicants correcting post-selection whose corrections do not result in an SEI change, the current eligibility status of applicants and the rate at which they appear on the recipient file; and
- the magnitude of change in a student's expected award resulting from the post-selection corrections.

It is important to analyze the PEC and random group using the above-mentioned measures for the following reasons:

1. Analysis of the number of applicants whose corrections do not result in SEI change indicates the degree to which the criterion groups are refined; that is, the degree to which the criterion group identify only those applicants whose award potential will change as a result of validation, and exclude those applicants likely to make corrections that do not impact on the student's award potential. It is assumed that it is most cost-effective to validate students in criterion groups that typically have the greatest SEI change. Analysis of this measure may suggest that certain criteria be excluded from the PEC because resources currently use in validation and processing corrections could better be used elsewhere.
2. Analysis of the relationship between the reason a student was chosen for validation (the PEC the student met) and the critical field(s) changed by the student post-selection should yield information useful in assessing whether the students' corrections are a function of the specific conditions identified by the PEC group or result from other factors.
3. Analyses of the magnitude of change in an applicant's expected award resulting from a post-selection correction should confirm the expected relationship between change in SEI and change in potential award level. The analyses will also indicate the potential savings resulting from the edits and validation for each PEC and the randomly-selected group.

4. Comparisons of the rate at which students appear on the recipient file, their current eligibility status, and the percentage of applicants submitting corrections not resulting in SEI change provide information which allows for the determination of whether certain edits are excessively restrictive and create unnecessary barriers for certain types of students.

The following sections analyze applicants selected for validation in terms of the above measures. Section 3.2 compares randomly-selected applicants with the total group of applicants selected because of PEC. Section 3.3 attempts to analyze the relative effectiveness of each PEC criterion by making within PEC group comparisons.

3.2 Overall Effectiveness of the PEC as Compared to Random Selection

This section assesses the overall effectiveness of the PEC relative to random selection as an indicator of misreporting. To be considered effective at this broad level, the PEC must, at a minimum, elicit a higher rate of applicants correcting and a greater resulting SEI change than the random group. These indicators provide some idea as to whether the concept of targeting specific groups of students for validation is appropriate.

The following table (Table 3.2) shows the percentage of PEC and random applicants making post-selection corrections, the resulting SEI change, and data on several other factors indicative of effectiveness. As can be seen in the table, approximately ten percent more PEC applicants than random applicants corrected any field post-selection. Although both applicant groups increased their SEI, the SEI change for the PEC applicants averaged almost 90 points more than for their randomly sellected counterparts. Furthermore, the ratio of PEC applicants raising their SEI to lowering their SEI is about twice that of the random group. The difference between the average amount SEI was raised and lowered was quite small for both groups, however, (+9 for the PEC group, and -9 for randomly selected applicants). The percentage of applicants appearing on the recipient file is also similar for both groups (approximately 57%). So, too, is the current eligibility status of the applicants.

TABLE 3.2: OVERALL COMPARISON OF EFFECTIVENESS OF RANDOM SELECTION VERSUS SELECTION ACCORDING TO PEC

	PEC Applicants	Randomly Selected Applicants
% of Applicants Correcting Any Field Post-Selection	38%	29%
Average SEI Change	128	37
Composite Change Index[1]	4,864	1,073
% of Applicants Correcting Any Field With No SEI Change	33.8%	64.1%
% of Applicants Correcting Critical Fields Post-Selection	32%	23%
Ratio of Applicants Raising SEI to Lowering SEI	3:1	1.7:1
Difference Between Average Amount Raised and Amount Lowered	9	-9
% of Applicants Correcting Critical Fields With No SEI Change	20%	29%
% of Applicants on Recipient File	56.2%	58.2%
Current Eligibility Status of all Applicants		
% Eligible	99.4%	99.6%
% Ineligible	.08%	.05%
% Rejected	.47%	.32%
Average $ Difference Between Selection and Current Payment Transaction[2]	-$20.2	-$3.1

[1] The Composite Change Index is an overall score of effectiveness derived by multiplying the rate of applicants correcting by their average SEI change.

[2] Based on all currently eligible applicants who are not on the recipient file.

Approximately 99.5 percent of all validation applicants--both PEC and random--are currently eligible. Between five and eight-tenths of one percent are currently ineligible, while less than one-half of a percent is rejected. Reflecting the significant difference in the percentage of applicants increasing their SEI and the magnitude by which the average SEI increased, PEC validation applicants have a much greater dollar difference between the average expected award at selection and that at payment than do randomly-selected applicants. The expected award declines by approximately twenty dollars between selection and the current transaction for PEC validation applicants, but by only about three dollars for randomly-selected validation applicants.

It can be concluded from the above data that, their totality, the pre-established criteria are more effective in identifying misreporters than the random selection group. This indicates that the basic concept of selective action strategies is sound. Applicants selected by this or a similar method will tend to have a higher percentage of applicants correcting and greater resulting SEI change than applicants selected for no specific reason. The data further indicate that, as would be expected, positive SEI change translates to smaller average expected awards. This means that applicants selected for validation based on the existing pre-established criteria will generally have a more significant reduction in potential award than randomly-selected students, resulting in fairly substantial savings. One additional piece of data that may also indicate that the PEC are better than random selection in identifying misreporters is that the percentage of applicants correcting whose correction do not result in SEI change is lower for the PEC group than for the random group, despite the fact that the overall correction rate is higher for applicants meeting PEC. One possible reason for this is that the PEC are indeed more sensitive to factors that result in SEI change. In sum, these analyses suggest that the development of criteria identifying misreporters is a concept that should be pursued.

3.3 Relative Effectiveness of PEC Subgroups

This section analyzes the relative effectiveness of the pre-established criteria. Although the PEC can generally be considered more effective than random selection in identifying students reporting inaccurate information, it is necessary to examine each PEC criterion separately to determine relative effectiveness when compared with other criteria and the random selection process. Previous studies have indicated that there is substantial variation in the degree of effectiveness within the various PEC subgroups. Certain PEC subgroups consistently score high when judged on effectiveness, while others score equal to or lower than the random group. A third group tends to be moderately effective. As with the previous analyses, effectiveness of the criteria is assessed primarily in terms of the rate of applicants correcting post-selection and the resulting average SEI change.

Table 3.3, on the following page, summarizes applicant corrections behavior and the resulting SEI change by PEC criteria and for the random applicant group. The table shows the number of applicants selected for validation for each criterion subgroup and group (A, B, C and D); the percentage of applicants correcting post-selection within each group; the average SEI change, the composite change index; and the percentage of applicants whose corrections do not result in SEI change. It should be noted that the table is based on corrections made to any field, not just to critical fields. Corrections counted include such things as address changes, and systems generated transactions, which do not result in SEI change. Therefore, the percentage of applicants whose corrections do not result in SEI change will be higher and the magnitude of the average SEI change will be lower than that found on Table 3.6, which analyzes correction made only to critical fields. Both tables are presented in this analyses to show that a significant number of the post-correction are made to other than critical fields.

3.3.1 Rate of Applicants Making Post-Selection Corrections

Table 3.3 clearly shows that there is substantial variation in the rate of applicants submitting post-selection corrections to any field

TABLE 3.3: SUMMARY OF VALIDATION APPLICANT POST-SELECTION CORRECTIONS BEHAVIOR AND RESULTING SEI CHANGE BY PEC AND RANDOM APPLICANTS[1]

CRITERION GROUP		NUMBER SELECTED	% OF APPLICANTS CORRECTING POST-SECTION	AVERAGE SEI CHANGE[2]	COMPOSITE CHANGE INDEX[3]	% OF APPLICANTS CORRECTING W/NO SEI CHANGE
PEC Applicant		116,348	38	128	4,864	33.8
PEC GROUP	A	137,618	38	138	5,244	35.3
	A-1	2,596	34	90	3,060	44.8
	A-2	68,800	38	165	3,420	38.0
	A-3	36,580	37	175	6,475	45.2
	A-4	18,130	29	53	1,537	77.9
	A-5	9,576	39	162	6,318	45.8
	A-6	44,948	42	233	9,786	23.5
	A-7	7,930	35	21	735	24.9
	A-8	6,986	34	103	3,502	26.5
	A-9	3,674	34	74	2,516	24.8
	A-10	21,832	49	132	6,468	12.7
	A-11	12,276	51	137	6,987	11.2
	A-12	2,720	30	138	4,140	54.8
	A-13	48	42	101	4,242	40.0
	A-14	7,120	49	113	5,537	18.9
	A-15	1,426	52	103	5,356	15.9
	A-16	438	29	- 36	-1,044	10.9
	A-17	7,006	30	125	3,750	31.1
	A-18	778	29	62	1,798	35.4
	A-19	3,746	27	131	3,537	35.4
	A-20	5,862	43	77	3,311	40.3
	A-21	63,342	40	117	4,680	31.8
Group	B	21,140	29	98	2,842	35.0
	B-1	686	33	- 22	- 726	[illegible]
	B-2	20,472	29	102	2,958	36.1
	B-3	44	36	233	8,388	12.8
	B-4	134	32	57	1,054	33.8
	B-5	410	39	93	3,627	20.3
	B-6	4,040	38	102	3,876	20.9
Group	C	13,274	29	111	3,219	34.5
	C-1	1,630	34	73	2,482	22.9
	C-2	11,674	28	117	3,276	36.8
	C-3	416	36	96	3,456	29.3
	C-4	3,542	35	116	4,060	24.4
Group	D	81,266	42	106	4,452	29.6
	D-1	15,406	51	75	3,825	23.4
	D-2	76,862	42	108	4,536	29.7
	D-3	11,002	56	76	4,256	17.0
Random		46,014	29	37	1,073	39.4

1/Corrections pertain to all fields, not only critical fields.

2/The average SEI change is derived by summing the SEI change for applicants raising and lowering their SEIs and dividing by the total number of applicants correcting, including those whose SEI did not change.

3/The Composite Change Index is an overall score of effectiveness derived by multiplying the percent of applicants correcting by the amount of their SEI change.

when compared by PEC group and subgroup. PEC group D had proportionately more applicants correcting post-selection than any other group (42%). PEC A followed with 38 percent of its applicants correcting. PEC groups C and B tied for the lowest percentage of applicants correcting post-selection (29%). The rate of corrections for PEC C and D group applicants equalled, but did not exceed, that of the random selection group.

The variation in the percentages of applicants correcting post-selection is more evident when PEC subgroups are examined. The table indicates that some PEC subgroups are extremely effective in terms of a high rate of applicants correcting, while others are less effective than random selection. The following subgroups have the highest and lowest rates of applicants correcting:

TABLE 3.4: PEC WITH HIGHEST CORRECTION RATE

PEC With the Highest Correction Rate		PEC With Lowest Correction Rate	
Applicant Subgroup	Magnitude of SEI Change	Applicant Subgroup	Magnitude of SEI Change
D-3	56%	B-14	22%
A-15	52%	A-19	27%
D-1	51%	C-2	28%
A-11	51%	A-4	29%
A-10	49%	A-16	29%
A-13	49%	A-18	29%
A-20	43%	B-2	29%
A-6	42%	A-12	30%
D-2	42%	A-17	30%
A-21	40%	A-17	30%

3.3.2 Average Effective SEI Changes

In contrast to the findings of prior analyses of correction rates, PEC groups A, B, C and D have similar average SEI changes. Group A has

an average change of 138 points, while groups C and D have average changes of 111 and 106 points, respectively. PEC group B trails, but insignificantly, with an average 98 point SEI change.

Further analyses of SEI change shows that there is considerable difference in the magnitude of SEI changes resulting from post-selection corrections when the changes are compared within PEC subgroups. The following table shows the PEC subgroups with the greatest and least SEI change.

TABLE 3.5: PEC WITH GREATEST AND LEAST SEI CHANGE

PEC With the Least SEI Change		PEC With the Greatest SEI Change	
Applicant Subgroup	Magnitude of SEI Change	Applicant Subgroup	Magnitude of SEI Change
A-7	+21	A-6	233
B-1	-22	B-3	233
A-16	-36	A-3	175
A-4	53	A-2	165
B-4	57	A-5	162
A-18	62	A-12	138
		A-11	137
		A-10	132

*Random group SEI change: 43 points.

As can be seen in the table, with the exception of criterion B-3 the PEC subgroups with the greatest SEI change are clustered in the A group. PEC subgroups and A-6 (income and taxes paid inconsistent), A-3 (unusual expenses inconsistent with income or greater than $5,000), A-2 (portions earned and AGI inconsistent), and A-5 (unreimbursed tuition greater than 35% of income) had post-selection correction which resulted in large positive SEI changes. The overall SEI change for applicants correcting from PEC subgroup B-3 was very large; however, the change was in the

negative direction. Applicants in PEC subgroups A-1 (missing information on household size or marital status), A-10 (no SS match and applicant verifies), A-11 (no SS match and applicant changes), A-12 (VA blank or zero and applicant verifies) and A-19 (combination of an A criterion with C-2) also had a high degree of SEI changes. The average change for these groups was 35 points.

The subgroups with the lowest magnitude of change were fairly evenly distributed between the A, B, and C groups. Both A-7 (financial data missing) and B-1 (large SEI change) had SEI change levels 6 points below that of the random group. Applicants in subgropus C-1 (large SEI change) and A-16 (combination of an A criterion and B-1) also had corrections which resulted in less SEI change than the randomly-selected groups.

3.3.3 Composite SEI Change

Another measure of the PEC's effectiveness in identifying potential misreporters is the composite SEI change index. This index is computed by multiplying the percentage of applicants correcting post-selection by the resulting average SEI change. It is based on the assumption that given the current system model which has one treatment for all applicants, the criteria determined most effective should be those that result in the highest overall yield to the Department of Education, both in terms of the percentage of applicants correcting and the magnitude of SEI changes. Using the composite change as an indicator, the following subgroups can be said to be most and least effective.

Therefore, the higher the number of applicants whose correction result in an SEI change, the more effective the criteria.

3.3.4 Percent of Applicants Correcting Post-Selection with no SEI Change

One final measure of effectiveness analyzed in Table 3.3 is the percentage of applicants correcting post-selection, whose corrections do not result in SEI change. Analyses of this measure provide further evidence of effectiveness by looking at the consistency by which the PEC

TABLE 3.6: MOST EFFECTIVE PEC AS DETERMINED BY COMPOSITE CHANGE INDEX[1]/

PEC Subgroup	% of Applicants Correcting Post Selection [2]/	Average SEI Change	Composite Change Index
A-6	42	233	9,786
B-3	36	233	8,388
A-11	51	137	6,987
A-3	37	175	6,475
A-10	49	132	6,468
A-5	39	162	6,318

*Composite score for random group: 1073

1/Subgroups are ranked in order of decreasing effectiveness when judged on composite score.

2/Represents individuals correcting any field post-selection, regardless of whether the correction resulted in an SEI change.

TABLE 3.7: LEAST EFFECTIVE PEC AS DETERMINED BY COMPOSITE CHANGE INDEX[1]/

PEC Subgroup	% of Applicants Correcting Post Selection [2]/	Average SEI Change	Composite Change Index
B-1	33	-22	- 726
A-7	35	21	- 735
A-16	29	-36	-1,044
A-4	29	53	1,537
A-18	29	62	1,798

*Composite score for random group: 1073

1/Subgroups are ranked in order of decreasing effectiveness when judged on composite score.

2/Represents individuals correcting any field post-selection, regardless of whether the correction resulted in an SEI change.

distinguishes between serious misreporters whose corrections will result in significant SEI change and other applicants. Ideally, the PEC should be very discriminating; to make validation most worthwhile only those students who have reported very inaccurate information on their application, and whose corrections will result in large SEI change should be selected for validation.

Data on Table 3.3 show that over one third of all validation applicants correcting post-selection submit corrections that do not result in an SEI change. The data also show that the rate of PEC applicants correcting without SEI change is only slightly less than the rate of randomly selected applicants correcting without change (33.8% and 39.4% respectively). Within the PEC subgroups, however, there is wide variation in the percentage of applicants correcting who have no SEI change. For example, over three quarters of all applicants correcting in PEC subgroup A-4, submitted corrections which had no impact on SEI, while only 7 percent of the applicants correcting in the B-1 subgroup had no SEI change. The table shows that the subgroups least consistent in identifying individuals whose corrections affect their award potential are (in order of greatest inconsistency): A-4 (77%), A-12 (55%), A-5 (45%), A-3 (45%), A-13 (40%). The most consistent PEC subgroups are: B-1 (7%), A-16 (11%), A-11 (11%), B-3 (12%), A-10 (13%), A-15 (16%), D-3 (17%) and A-14 (19%).

TABLE 3.8: MOST AND LEAST CONSISTENT PEC GROUPS

PEC Subgroups Most Consistent in Identifying Students With SEI Change		PEC Subgroups Least Consistent in Identifying Students With SEI Change
B-1	A-15	A-4
A-16	D-3	A-12
A-11	A-14	A-5
B-3		A-3
A-10		A-13

It should be noted that there are several reasons why the rate of corrections not resulting in SEI change seems to be high. First, Table 3.3 is based on corrections made to any field, and as such, includes corrections (such as address changes and systems generated transactions) that do not affect a students SEI. A certain percentage of non-SEI change corrections can, therefore, be attributed to these factors. However, our analyses suggest address changes and the like make up only a small percentage of the total corrections. Table 3.4 on the following page, shows that almost 85 percent of the post-selection corrections are made to critical fields and therefore have the possibility of causing SEI change. The second reason that the correction without SEI change rate may be high is that the applicants correcting have negative SEIs, whicha re set to zero by he eligibility formuala, that changes, even fairly major ones, do not raise their SEI above zero. Finally, the rate may be high because the corrections made by the applicant are so minor that they don't alter the student's SEI, regardless of the initial SEI.

3.3.5 Change to Critical Fields

Table 3.9 is included in this report because it can be argued that analyses of corrections made to critical fields tend to provide more precise information on the number and impact of corrections than data on corrections made to any field. Precision is increased because the corrections rate and average SEI change scores are not diluted by including in the population those applicants making corrections that by definition cannot result in SEI change.

Despite the differences in the way the corrections rate and average SEI change were computed in Tables 3.3 and 3.9, the PEC subgroups that are most and least effective when judged by the composite index (the score that takes into account both the rate of applicants correcting and the average SEI change) are similar. The subgroups considered most effective when analyzed according to critical field coorections are subgroups A-3, A-5, A-6, A-11, and B-3; the subgroups which continue to be least effective are A-4, A-7, A-16, A-15, and B-1.

BEST COPY AVAILABLE

TABLE 3.9: POST-SECTION CORRECTION RATE TO CRITICAL FIELDS, AVERAGE SEI CHANGE AND COMPOSITE SEI CHANGE BY VALIDATION APPLICANT SUBGROUP

Applicant Subgroup	Number Selected	% of Applicants Correcting Post-Selection 1/	Average Effective SEI Change	Composite SEI Change Index 2/	% of Applicants Raising SEI 4/	Average Amount Raised	% of Applicants Lowering SEI 4/	Average Amount Lowered	Ratio: # Raised to # Lowered 3/	Average Diff. Between Amount Raised and Lowered	% of Applicants Correcting with No SEI Change
PEC Applicants	166,348	32	144	4608	55	349	19	-254	3:1	95	26
PEC Group A	137,618	32	155	4960	54	349	18	-340	4:1	9	28
A-1	2,596	29	105	3045	43	409	22	-316	2:1	93	35
A-2	68,800	33	185	6105	53	444	17	-291	3:1	153	30
A-3	16,580	33	199	6567	48	498	14	-306	3.5:1	192	38
A-4	18,130	20	69	1380	20	555	9	-439	2:1	116	71
A-5	9,576	34	185	6290	47	486	15	-288	3:1	198	38
A-6	44,948	39	255	9945	67	449	16	-296	4:1	153	17
A-7	7,930	31	23	713	51	279	34	-345	1.5:1	66	15
A-8	8,988	30	114	3420	50	334	24	-332	2.5:1	2	18
A-9	3,674	31	83	2573	56	253	27	-219	2:1	34	17
A-10	21,832	35	137	4795	71	251	21	-196	3.5:1	55	8
A-11	12,276	37	144	5328	73	240	19	-166	4:1	74	8
A-12	2,720	24	158	3792	45	436	8	440	5.5:1	4	47
A-13	48	33	126	4158	50	376	25	-249	2:1	127	25
A-14	7,120	41	117	4797	61	268	24	-190	-1	78	15
A-15	1,426	50	105	5250	62	259	25	-222	3:1	37	23
A-16	438	27	- 39	-1053	59	165	37	-367	1.5:1	-202	4
A-17	7,006	25	143	3575	57	333	23	-208	2.5:1	125	10
A-18	778	25	70	1750	63	230	22	-332	3:1	-102	15
A-19	3,746	22	151	3322	56	331	19	-166	3:1	165	25
A-20	5,862	39	85	3315	51	225	14	-213	3.5:1	12	35
A-21	63,342	16	128	4608	54	348	21	-294	2.5:1	54	25
PEC Group B	21,140	24	113	2712	52	221	36	-363	1.5:1	-142	12
B-1	686	31	- 23	- 713	58	178	39	-322	1.5:1	-144	3
B-2	20,472	24	119	2856	52	337	24	-219	2:1	98	24
B-3	44	36	233	8388	88	266	0	0	88:0	266	0
B-4	134	19	66	1254	46	388	31	-367	1.5:1	21	23
B-5	410	34	107	3638	71	197	20	-162	3.5:1	35	9
B-6	4,040	35	110	3850	58	300	27	-259	2:1	49	15
PEC Group C	13,274	23	128	2944	55	307	22	-191	2.5:1	116	23
C-1	1,630	30	79	2370	66	196	26	-254	2.5:1	- 58	14
C-2	11,674	22	138	4140	53	333	22	-180	2.5:1		25
C-3	416	31	110	3300	63	198	19	-72	3.5:1	126	18
C-4	1,102	31	124	3844	61	289	23	-225	2.5:1	64	16
PEC Group D	81,266	38	116	4408	56	301	21	-254	2.5:1	47	23
D-1	15,406	48	79	3792	61	173	20	-129	3:1	44	19
D-2	76,862	38	117	4446	56	305	21	-256	2.5:1	49	23
D-3	11,002	54	79	4266	64	167	22	-125	3:1	42	14
Random	46,014	21	43	909	45	286	26	-245	1.5:1	9	29

1/PEC subgroup percentages are based on unduplicated applicant totals within, but not between, groups.

2/The Composite SEI Change Index is derived by multiplying the percentage of applicants making post selection corrections to critical fields by the average effective SEI change.

3/Ratios are rounded to nearest half of one percent.

4/The last three percent columns (% of Applicants Raising SEI, % of Applicants Lowering SEI + % of Applicants Correcting with No SEI Change) add up to 100% and all components of the first percent column (% of Applicants Correcting Post Selection)

Table 3.9 is included in this report for two other reasons. First, it allows one to probe deeper and examine not only the average SEI change, but also the number of applicants missing and lowering thier SEI. Second, it shows that a large number of applicants correcting without SEI change are making their corrections to critical fields. This has significant implication regarding the reliability of some PEC subgroups. The high rate of applicants correcting with no SEI change cannot be excused as resulting merely from address changes and similar factors.

Analysis of the percentage of applicants and their SEI show that there is a substantial difference in the ratios of those increasing and decreasing their award potential by PEC group and subgroup. Overall the number of PEC applicants correcting to their disadvantage outweighs the number correcting to their advantage by almost three to one. Random applicants also submit corrections that result in more applicants increasing their SEI than decreasing it; however, the ratio of random applicants raising to lowering their SEI is significantly lower than that for PEC applicants. Comparison of the four PEC groups shows that applicants in PEC group A tend to, on the average, increase their SEI more often than applicants in the other PEC groups. The ratios of applicants increasing their SEI to decreasing it, is as follows for each PEC group: A (3:1); B (1.5:1); C (2.5:1); D (2.5:1).

The variation in the ratios of applicants correcting to their disadvantage and advantage within PEC subgroups are greatest in groups A and B. In group A, the proportion of applicants increasing their SEI up to those decreasing it ranges from 1.2:1 for subgroup A-7 to 5.5:1 for subgroup A-12. In group B the ratios range from 1.5:1 to one subgroup where no applicants decreased their SEI (subgroup B-3). As mentioned previously, the value of the numerical averages presented for PEC subgroups B-1, B-3, B-4 and B-5 is somewhat questionable because of the small number of subjects in the study population.

Analysis of the difference between the magnitude of SEI change of the applicants correcting to their disadvantage and those of applicants

correcting to their advantage (in contrast to the number of students raising and lowering their SEI), indicates that the average SEI change of applicants decreasing their award potential is only slightly higher than that of PEC applicants decreasing their potential. The average amount by which random group applicants increase their award potential, on the other hand, is slightly more than the amount of those decreasing their potential award value.

As shown in the analyses of other effectiveness measures, there is much variation in the behavior of applicants within the PEC subgroups. This variation is evident in the analysis of the average difference between the amount SEI's were raised and lowered, and the direction of the net change. For example, in 18 of the total 34 PEC subgroups, the magnitude of change in one direction outweighs the magnitude of SEI change in the opposite direction by more than 100 points. However, in 12 groups the net difference is less than 50 points. Similarly, the net difference between the average amount the SEI was raised and lowered, was positive for 29 of the PEC subgroups, but negative for the randomly-selected group and other PEC subgroups.

Finally, as mentioned previously, this table shows that most of the corrections not resulting in SEI change are being made to critical fields. (In addition, when corrections to critical fields are analyzed, PEC subgroups continue to have exceptionally high rates of applicants with no SEI change.) This indicates that current criteria are not adequately refined to identify misreporters, and that other methods of selecting applicants should be studied.

3.3.6 Relationship of Corrections Behavior of PEC Validation Applicants to "Suspect" Fields

The majority of the A and all of the D PEC criteria are based on the assumption that potential misreporters can be identified solely because of the absence of certain information on their application or the presence of inconsistent, illogical or other questionable application data. It is assumed that either the illogical or inconsistent data are incorrect, and when validated will be corrected, or that other related data are inaccurate and will also be corrected as a result of validation.

TABLE 3.10: PERCENTAGES OF VALIDATION APPLICANTS WITH POST-SELECTION CORRECTIONS TO SUSPECT FIELDS BY PEC GROUP[1/]

PEC SUB-GROUPS	% OF VALIDATION APPLICANTS CORRECTING FIELD — CRITICAL FIELDS													
	ADJUSTED GROSS INCOME	NONTAXABLE INCOME [2/]	TAXES PAID	MODEL	HOUSEHOLD SIZE	POST HIGH EDUCATION	NET ASSETS	MARITAL STATUS	VETERAN'S BENEFITS	UNREIMBURSED TUITION	APPLICANT RESOURCES	UNUSUAL EXPENSES	TAX FILING STATUS	PORTION EARNED
A-1	18	11	12	2	7	3	3	2	2	0	1	4	6	12
A-2	21	9	15	1	6	4	3	1	1	1	1	9	6	33
A-3	21	7	13	1	5	4	3	1	1	1	1	37	5	12
A-4	8	16	3	2	5	2	2	3	2	0	1	4	3	7
A-5	20	9	13	0	5	5	3	1	1	6	1	13	5	11
A-6	24	8	24	0	6	5	3	1	1	1	1	10	6	13
A-7	13	12	12	0	6	5	4	1	1	1	1	2	6	8
A-8	15	14	12	1	6	4	3	1	1	1	1	3	4	9
A-9	14	12	12	0	6	4	3	1	1	0	5	2	3	8
A-10	17	51	13	1	6	4	3	1	3	0	1	3	4	9
A-11	18	54	14	1	6	4	3	1	3	0	1	2	4	10
A-12	10	13	6	1	4	2	1	2	12	1	0	3	3	9
A-13	13	17	8	0	0	0	0	0	17	0	0	4	0	8
A-14	15	43	11	1	6	4	3	1	14	0	1	4	3	8
A-15	18	43	13	1	6	3	3	2	14	0	1	5	4	9
A-16	11	7	13	0	5	3	4	0	0	0	1	1	2	5
A-17	11	10	9	0	5	4	3	1	1	1	1	3	3	7
A-18	12	11	10	1	4	2	2	1	2	0	2	1	2	6
A-19	9	11	8	0	5	3	2	1	1	1	0	3	2	6
A-20	19	13	15	0	9	17	3	1	1	2	0	6	5	12
A-21	23	11	19	1	6	4	3	1	1	2	1	6	3	15
B-1	18	6	19	0	4	4	4	0	1	1	1	1	6	7
B-2	12	11	9	1	5	3	3	1	1	1	1	2	3	7
B-3	23	5	14	0	0	9	0	0	0	0	5	0	5	5
B-4	3	6	7	1	4	0	3	3	3	0	0	0	7	6
B-5	17	4	18	0	6	13	5	0	0	1	0	2	3	10
B-6	24	8	19	0	5	4	4	1	0	1	1	3	10	13
C-1	17	9	15	0	5	3	3	0	1	0	1	1	5	9
C-2	10	12	8	1	5	3	2	1	1	1	0	2	3	6
C-3	12	10	13	0	8	10	2	1	0	1	1	1	4	7
C-4	19	9	17	1	5	3	4	1	0	0	1	2	7	11
D-1	36	8	31	0	9	15	4	1	1	1	1	4	11	18
D-2	26	10	22	1	6	5	3	1	1	1	1	5	10	16
D-3	44	6	39	0	10	16	4	1	0	1	1	3	14	21

[1/] Shaded areas denote critical fields considered "suspect" for each PEC group. Bold face numerals signify areas of greatest change.

[2/] Social security benefits are included as part of Nontaxable Income.

In assessing the overall effectiveness of the criterion groups, it is beneficial to analyze this relationship between the reason a student was chosen for validation (the PEC criteria met) and the critical fields changed by the student post-selection. Although one cannot assume that there will always be a direct relationship between PEC subgroup applicants and suspect fields changed, one can assume that the stronger the relationship, the more sensitive the subgroup in identifying misreporters.

Table 3.10 on the preceding page shows that there is some correlation, albeit a small one, between the reason for selection and the suspect fields corrected for most A and D PEC subgroups. This suggests that there is a large random element in misreporting. The PEC subgroups with the highest rate of corrections to suspect fields, hence the strongest correlation, are A-2, A-3, A-6, A-7, A-10, A-11, A-13, A-14, A-15, and D-20. It is interesting to note that these subgroups are among the ones considered most effective when evaluated in terms of the percentage of applicants correcting and the magnitude of the resulting SEI change. These groups are also among the most consistent in eliciting corrections that result in SEI changes. PEC subgroups A-1, A-4, and A-9 have the most tenuous relationship between the reason for selection and corrections to corresponding suspect fields. In most cases these subgroups have also performed relatively poorly when evaluated on the other efficacy measures.

More important than the data presented above, Table 3.10 clearly shows that AGI and Taxes Paid are the fields corrected most often regardless of the PEC criteria meet. Portions Earned and Nontaxable Income are also corrected at a much higher rate than other fields, but at a rate slightly lower than AGI and Taxes Paid. It is difficult to determine why these fields are corrected at such a high rate; however, one reason may be that tax and income information is relatively easy to verify and therefore may be subjected most rigorously to validation.

Analysis of PEC subgroup A-5 yielded some particularly interesting findings. This group has a relatively high percentage of applicants

correcting and a relatively high degree of SEI change, indicating that it is a fairly effective criterion. Yet, it does not show a high percentage of applicants correcting the unreimbursed tuition field post-selection. Instead, a relatively large percent of the corrections made by A-5 applicants are made to income-related fields, specifically AGI and Taxes Paid. This suggests that the ratio of the percent of unreimbursed tuition to income currently used for the criteria is fairly refined in identifying individuals who under report their income.

3.3.7 Rate of Applicants Appearing on the Recipient File and Their Current Eligibility Status

We have determined the rate at which PEC and randomly-selected applicants do not appear on the recipient file and the current eligibility status of these applicants. These factors, although not clear indicators of effectiveness, allow us to generate some hypotheses on how PEC and randomly-selected validation applicants negotiate the processing system edits and validation.

Table 3.11, on the following page, shows the percentage of validation applicants not appearing on the recipient field as of May, 1980 and the percentage of these applicants who are currently eligible. The table indicates that there are only slight differences in the percentage of applicants expected to be paid and the percent currently eligible, for the total PEC applicant and random applicant groups. Except for PEC group C, there is also very little difference in these scores when compared across PEC groups. For example, the percentage of applicants expected to be paid averages 43.8 percent for the A group, 42.0 percent for the B group, and 43.4 percent for the D group. The percentage of applicants on the recipient file is also similar for these three groups (A = 94.1%; B = 94.8%; D = 93.5%). On the other hand, PEC group C has quite different averages. Both the total percent of applicants not expected to be paid and the rate of currently eligible applicants within this category are much lower than for the other PEC groups.

TABLE 3.11: PERCENTAGE OF VALIDATION APPLICANTS NOT EXPECTED TO BE PAID AND THEIR CURRENT ELIGIBILITY STATUS

Criterion Group	% Not Expected To Be Paid	% Not Expected To Be Paid But Currently Eligible	Criterion Group	% Not Expected To Be Paid	% Not Expected To Be Paid But Currently Eligible
PEC Applicants	43.8	94.2	B-2	42.0	95.0
A	44.7	93.9	B-3	50.0	90.9
A-1	59.2	95.2	B-4	34.3	100.0
A-2	44.2	93.4	B-5	40.5	92.8
A-3	42.6	93.1	B-6	43.0	92.3
A-4	49.7	96.1	C	35.5	97.4
A-5	42.1	93.2	C-1	33.6	96.7
A-6	45.2	91.1	C-2	35.8	97.5
A-7	40.6	93.6	C-3	27.4	100.0
A-8	41.2	94.7	C-4	36.2	95.9
A-9	38.9	93.0	D	43.4	93.5
A-10	44.6	94.4	D-1	40.1	94.7
A-11	44.2	94.3	D-2	43.6	93.4
A-12	64.0	96.3	D-3	40.2	94.2
A-13	66.7	93.8			
A-14	39.0	93.44	Random	41.8	96.5
A-15	36.6	94.6			
A-16	39.7	91.9			
A-17	41.3	93.9			
A-18	34.4	97.8			
A-19	37.4	97.9			
A-20	41.7	95.0			
A-21	44.5	93.2			
B	42.0	94.8			
B-1	40.2	89.1			

There is more variation in scores within PEC subgroups, although even within these groups the variation is not great. The subgroups with the highest percentage of applicants not on the recipient file are: A-1 (59.2%), A-4 (49.7%), A-6 (45.2%), A-12 (64.0%), A-13 (66.7%) and B-3 (50.0%). Only PEC subgroup C-3 has a percentage of applicants not on the recipient file that is substantially lower than other groups (27.4%). The subgroups with the greatest percentage of not expected to be paid currently eligible applicants are A-4 (96.1%), A-12 (96.8%), A-18 (97.8%), A-19 (97.9%), B-4 (100%), C-1 (96-7%), C-2 (97.5%) and C-3 (100%). PEC subgroups A-6, A-10, A-11, A-13, A-14, A-15, A-18, B-3, and C-4 also have the greatest dollar difference. These groups also correspond to the groups with the highest SEI changes. In all but subgroup A-13, the difference results in savings to OSFA. It appears that applicants in A-13 tend to underestimate their financial need and that validation assists these applicants in obtaining greater awards.

PEC subgroups A-1, A-4, A-7, A-9, A-16, and B-1 have the lowest dollar difference between expected payment at selection and at the current transaction. These subgroups had an average dollar savings of only $9,069 per subgroup. This is not suprising, however, since these subgroups have performed poorly on almost every other effectiveness measure.

Analysis of these figures with the figures on the percentage of applicants correcting whose corrections do not result in any SEI change Table 3.11 suggests that there may be some correlation between these numbers. For example, the subgroups with the highest percentage of applicants not on the recipient file (A-1, A-4, A-12, A-13, and B-3) tend to have either very high or very low percentage of applicants correcting without resulting SEI change. The following table summarizes these relationships. It can be hypothesized that high not expected to be paid scores combined with a high percentage of applicants correcting without SEI change indicates unnecessary barriers placed in the way of applicants in the groups having these scores. The edits may also be restrictive and

TABLE 3.12: HYPOTHESIZED RELATIONSHIP BETWEEN APPLIANTS NOT ON THE RECIPIENT FILE AND APPLICANTS CORRECTING WITHOUT SEI CHANGE

PEC Subgroups	% Not Expected to be Paid[1/]	% of Applicants Correcting Without SEI Change[2/]
A-1	59.2	44.6
A-4	49.7	77.9
A-12	64.0	54.8
A-13	66.7	40.0
B-3	50.0	12.5

1/Average not expected to be paid score for PEC applicants is 43.8 percent.

2/Average percentage of applicants correcting with no SEI change for PEC applicants is 33.8 percent (see Table 3.3).

may be eliciting a high rate of unnecessary corrections. Similarly, a high percentage of not expected to be paid applicants, combined with a low percentage of applicants correcting with no SEI change, could indicate subgroups effectiveness in identifying actual misreporters (B-3). The high percentage of applicants correcting whose corrections results in an average increase SEI suggests that the criterion is fairly refined in identifying individuals who are serious misreporters. The fact that a relatively small percentage of these applicants appear on the recipient file could indicate that either (1) the misreporting was detected during validation and the resulting corrections rendered the student ineligible or (2) that the threat of validation frightened many misreporters out of the system.

Examination of the percentage of currently eligible applicants who are not expected to be paid lends credence to these hypotheses. As can be seen in Table 3.1 , both very high and very low percentages of currently eligible applicants not on the recipient file correlate with

TABLE 3.13: ESTIMATED DOLLAR SAVINGS FOR CURRENTLY ELIGIBLE VALIDATION APPLICANTS NOT ON THE RECIPIENT FILE

Criterion Group	Average $ Differences Between Selection & Current Transaction1/	# of Applicants Currently Eligible	Estimated Savings2/
PEC Applicants			
PEC			
Group			

very high percentages of the total applicants not expected to be paid. Examination of subgroups of particular interest above, (A-1, A-4, A-12, A-13, B-3), shows that the high percentage of currently eligible applicants in groups A-1, A-4, and A-12 supports the hypotheses that applicants in these groups are being unnecessarily hindered by the current edit and validation systems. Conversely, the low percentage of currently eligible nonrecipient applicants found in D-3 suggests that these applicants may be legitimately excluded from receiving awards.

3.3.8 Expected Dollar Savings

The final measure for analysis in the study of the effectiveness of the PEC is estimated dollar savings. As can be seen in Table 3.13, significantly greater savings can be expected by selecting validation students based on PEC criterion instead of randomly. Students selected for validation by PEC but not currently on the recipient file, have SEI changes which may result in substantially greater savings for the PEC than the random group.

Within the PEC groups, PEC groups A and D had the greatest average dollar difference between the amount of expected payment at selection and the amount on the most current transaction, thus indicating the two most promising large groups for selection. As would be expected, these groups are also the ones with the greatest SEI changes (see Section 3.2).

3.4 Summary of Findings

Our analyses of the various measures of effectiveness for the pre-established criteria indicate that there is wide variation in the effectiveness of the group and subgroups in identifying students likely to misreport. The analyses indicate that in general, however, the concept of selecting students for validation according to specific characteristics that tend to be indicative of misreporting, i.e., inconsistent taxes and income information,is sound. This is evidenced by the fact that the overall average scores for PEC applicants are higher than for random applicants. Analyses of the four PEC group A, B, C, and D suggest that only groups A and D consistently identify a relatively

high percentage of misreporters. Groups A and D have overall composite change indices that are greater than for the indices of B and C groups PEC group A (4864); B (2842); C (3219); and D (4452). In contrast to the B and C groups which identify students for validation based on the students current and previous questionable transactional history, group A and D are based on current application data such as veteran's status, inconsistencies on income and tax information. This has implications for information stored on all applicants file suggests that current data on a students application form may be sufficient to predict misreporting.

There is also much variation in the efficacy of the PEC subgroup. When judged on a composite score taking into account both post-selection corrections rate and average SEI change, criteria A-3, A-5, A-6, A-10, A-11, A-14, A-15 and B-3 can be considered most effective and criteria A-1, A-4, A-7, A-9, A-16, A-18, and B-4 least effective.

Seven of the above mentioned effectiveness criteria are particularly interesting because they had high scores on the gross measures of effectiveness (percent of applicants correcting and average SEI change) and also tended to be quite refined in identifying applicants whose corrections wre not inconsequential, but instead resulted in SEI change. In addition, applicants in these groups tended to have a lower rate of currently eligible applicants considered not expected to be paid, possibly indicating that these students are in fact, misreporting and are either being detected at validation and forced to make corrections (possibly ones excluding them from the system) or else are being frightened out of the system by the prospect of validation. These seven pre-established criteria and their scores in the efficacy measures are seen in Table 3.14.

Similarly, there are at least two subgroups that neither elect large net composite SEI change nor are particularly refined. The subgroups are A-1 and A-4 (see Table 3.14).

BEST COPY AVAILABLE

Applicant Subgroup	% Correcting Post-Selection	Average SEI Change	Composite Change Index	% Correction With No SEI Change	% Not Expected to be Paid but Currently Eligible
A-5 (unreimbursed tuition greater than 35% of income)	39%	162	6318	45.5%	93.2%
A-6 (AGI and Taxes Paid Inconsistent)	42%	233	9786	23.5%	91.1%
A-10 (No SSA match and applicant verifies)	49%	132	6468	12.7%	94.4%
A-11 (No SSA match and applicant changes)	51%	137	6987	11.2%	94.3%
A-14 (VA 0 $131, applicant verifies)	49%	133	5537	18.9%	93.4%
A-15 (VA 0 $131, applicant changes)	52%	103	5356	15.9%	93.4%
B-3 (combination B-1 and C-1)	36%	233	8388	12.5%	90.0%

TABLE 3.14: THE LEAST EFFECTIVE AND REFINED PRE-ESTABLISHED CRITERIA

Applicant Subgroup	% Correcting Post-Selection	Average SEI Change	Composite Change Index	% Correction With No SEI Change	% Not Expected to be Paid but Currently Eligible
A-1 (Household Size or Marital Status Missing)	34	90	3060	44.6%	95.2%
A-4 (now Income Applicants)	29	53	1537	77.9%	96.1%

4

VALIDATION AND THE INTERACTION OF VALIDATION AND EDITS

Each year financial aid administrators at institutions of higher education across the country validate approximately 10 percent of the Basic Grant applicant popoulation. As of May 1980, 212,362 individuals had been selected for validation and had presumably documented their financial status to officials at the school they plan to attend. Approximately 78 percent of the applicants were chosen for validation because they meet one or more of the pre-established criteria; the other 22 percent were selected randomly.[1/]

This chapter analyzes the impact of validation[2/] on applicants post-selection/eligibility corrections behavior and SEI change, and the relative impact of the processing system edits and validation on the correct award of Basic Grants. The first part of this Chapter (Sections 4.1-4.3) examines the pre- and post-selection/eligibility corrections history of validation and nonvalidation applicants to determine whether the applicant groups are comparable, and to ascertain which pre-established criteria are most affected by validation. The primary analytic measures used in this section are (1) the rate of pre- and post-selection/eligibility corrections, (2) the resulting average SEI change,

---

1/See page for discussion of meaning of random selection.

2/More accurately, this chapter analyzes the impact of being selected for validation since there is no way of knowing whether validation actually took place.

(3) the percentage of applicants correcting post-selection without changing their SEI, and (4) the percentage of applicants expected and not expected to be paid and their current eligibility status. The second part of this chapter (Section 4.4) examines the interaction of the edits and validation to determine which process is better in assuring that applicants report correct application information. The validity of critical data elements at various points in the edits and validation process is the primary measure for analysis.

4.1 Impact of Validation-Comparison of Pre-and-Post Selection/ Eligibility Change

The following table (Table 4.1) shows the pre- and post-selection/ eligibility corrections behavior and resulting average SEI change for validation and nonvalidation applicants. The table is divided into all PEC applicants, PEC group and subgroup applicants, and those randomly selected or not meeting the PEC.

Analysis of all validation and nonvalidation PEC group A applicants shows that the rate of applicants correcting pre-selection/eligibility and the resulting SEI change is almost identical for both groups.[1/] Seventy-four percent of all PEC validation applicants corrected pre-selection and had corrections which resulted in an average 99 point SEI change; seventy-three percent of the nonvalidation PEC group A applicants corrected pre-eligibility with a resulting 98 point average SEI change.

There is, however, very substantial difference in the post-selection/ eligibility behavior of these two applicant groups. Almost five times more PEC group A validation applicants made post-selection corrections than nonvalidation group A applicants (32% of validation applicants

---

[2/] Before preceding further, it should be mentioned that although comparison of all PEC validation applicants with all nonvalidation PEC applicants is the logical first measure for analysis, this comparison is not possible. As mentioned in Chapter 1, nonvalidation applicants meeting PEC groups B, C. and D were not flagged by the central processor are therefore not correctly identified in the data base and cannot be compared.

BEST COPY AVAILABLE

TABLE 4:1: PRE -AND POST- SELECTION (ELIGIBILITY) CORRECTION BEHAVIOR OF VALIDATION AND NONVALIDATION APPLICANTS

	VALIDATION APPLICANTS						NONVALIDATION APPLICANTS					
Applicant Subgroups	Number of Applicants Selected	% of Validation Population	Percentage of Subgroup Applicants With Pre-Selection Field Changes	Average SFI Change	Percentage of Subgroup With Applicants POST-SELECTION Field Changes	Average SFI Change	Number of Applicants Selected	% of Validation Population	Percentage of Subgroup Applicants With Pre-Selection Field Changes	Average SFI Change	Percentage of Subgroup Applicants With POST-SELECTION Field Changes	Average SFI Change
PFT												
Applicants	166,148	78	71	-64	12	144	506,912	19	73	98	6	88
PFT												
Group A	137,618	65	74	99	32	155	506,912	19	73	98	6	88
A-1	2,596	1	88	186	29	105	10,584	1	84	236	6	0
A-2	68,880	32	81	217	33	185	124,376	5	78	148	6	22
A-3	36,580	17	77	283	33	199	41,314	2	70	180	6	61
A-4	10,130	9	55	36	20	69	112,924	4	48	12	5	33
A-5	9,576	5	82	292	34	185	9,296	1	88	175	6	85
A-6	44,948	21	78	290	39	255	127,764	5	85	195	6	80
A-7	7,930	4	97	170	31	23	27,552	1	94	105	6	40
A-8	8,988	4	91	63	30	114	79,240	3	88	36	6	15
A-9	3,674	2	98	641	31	83	1,078	1	66	371	36	143
A-10	21,832	10	60	92	35	137	51,310	2	82	230	5	8
A-11	12,276	6	85	100	37	144	2,086	1	86	114	11	44
A-12	2,720	1	42	3	24	158	5,390	1	85	127	7	15
A-13	48	1	83	47	33	126						
A-14	1,120	3	68	68	41	117	17,906	1	66	61	8	89
A-15	1,476	1	93	81	50	105						
A-16	488	1	98	168	27	-39						
A-17	1,006	3	95	-450	25	143						
A-18	778	1	97	-132	25	70						
A-19	1,746	2	92	280	22	151						
A-20	5,862	3	75	64	39	85						
A-21	63,142	1	79	160	16	128						
PFT												
Group B	21,140	10	95	818	24	113						
B-1	686	0	90	10	31	-23						
B-2	20,472	10	94	846	24	119						
B-3												
B-4	134	1	97	-702	19	66						
B-5	410	1	96	-713	34	107						
B-6	4,040	2	97	596	35	110						
PFT												
Group C	13,274	6	94	520	23	128						
C-1	1,630	1	97	215	30	79						
C-2	11,674	5	94	571	22	138						
C-3	416	1	97	-579	31	110						
C-4	1,542	2	98	467	31	124						
PFT												
Group D	81,766	38	73	81	38	116						
D-1	15,306	7	41	-9	48	79						
D-2	76,862	36	73	89	38	117						
D-3	1,102	1	28	57	54	79						
Random	46,014	22	25	44	23	43	2,116,912	81	6	26	4	47
TOTAL	212,162	100	64	55	30	127	2,623,824	100	19	65	4	44

and 6% of the nonvalidation applicants corrected post-selection/eligibility). The resulting average SEI change is also significantly different for both groups. The average SEI change for group A validation applicants is 155 points; the average change for group nonvalidation applicants is 38 points. Thus, the average post selection SEI change for validation applicants is over three times greater than for nonvalidation applicants.

Comparison of the corrections rate and SEI change for validation and nonvalidation PEC A applicants strongly indicates that validation is having an impact on applicants' post-selection/eligibility corrections behavior and resulting SEI change. Almost 20 percent more validation applicants than nonvalidation applicants corrected post-selection/eligibility, and their corrections resulted in an average 120 point greater SEI change.

Comparison of randomly selected validation applicants and nonvalidation applicants not meeting PEC showed that these group behaved differently than the groups mentioned above. Randomly selected validation applicants corrected pre-selection almost five time more often than nonvalidation applicants (25% and 6% respectively). In addition, although the magnitude of the SEI change was not that different for the two groups (44 points for validation applicants; 26 points for nonvalidation applicants), the direction of SEI change was dissimilar. Ramdomly selected validation applicants generally increased their SEIs pre-selection while on nonvalidation applicants decreased their SEIs.

Analyses of the rate of validation and nonvalidation applicants correcting pre-selection/eligibility by specific criterion subgroups shows that there is very little difference in the rate at which one validation subgroup corrects and the rate at which its companion nonvalidation subgroup corrects. The only exceptions to this are subgroups A-9 and A-12. Proportionately, one third more validation applicants than nonvalidation applicants corrected pre-selection in subgroup A-9, while fifty percent fewer validation applicants corrected pre-selection in subgroup A-12.

There is greater variation in the average SEI change for the validation and nonvalidation subgroups. Although, all but three subgroups have net SEI changes in the same direction (the exceptions are subgroups A-4, A-7, and A-12) the difference in the magnitude of the changes varies significantly. For example, there is a 270 point difference between the average SEI change for validation and nonvalidation A-9 subgroup applicants, but only a 14 point difference in the SEI change for A-11 applicants.

Both validation and nonvalidation applicants in subgroups A-2, A-3, A-5, A-6, A-7, A-10, A-11, A-14 made pre-selection/eligibility corrections which resulted in net positive SEI change. Validation and nonvalidation applicants in subgroups A-1 and A-9, however, made changes which resulted in a net decrease in SEI. Validation applicants in subgroups A-4, A-8, and A-12 decreased their SEIs pre-selection, while their nonvalidation counterparts initially increased their eligibility index.

Analyses of post-selection corrections and average SEI changes for validation and nonvalidation applicants in PEC A subgroup showed that with the exception of four subgroups (A-4, A-8, A-11,-12), between five and six times more validation applicants corrected post-selection than nonvalidation applicants. The SEI change resulting from these corrections differs within the subgroups. The difference between the average validation SEI change was greater than 150 points for subgroups A-2 and A-6, between 100 and 150 points for A-1, A-3, A-5, A-10 and A-12, and less than 100 points for subgroups A-4, A-8 and A-9. The SEI change for validation applicants in subgroup A-4 was only slightly above the change for comparable nonvalidation applicants, while the SEI change for validation applicants in subgroup A-7 and A-9 was less than that for similar nonvalidation applicants.

4.2: Impact of Validation - Percent of Applicants Correcting Post-Selection with No SEI Change

Analyses of miscellaneous data elements brought to light one element which seems to indicate that validation helps applicants make appropriate

decisions on when it is necessary to correct their application information. Table 4.2 shows that although over four times as many validation as nonvalidation PEC applicants corrected (38% to 9%, respectively), the rate of validation PEC applicants with no resulting SEI change is only about half that of the nonvalidation group (33.8% to 64.1%). The figures for the random group are similar. This suggests that the realtively small percentage of nonvalidation applicants who correct are making inconsequential, and most likely random corrections. The figures for the validation group, on the other hand, indicates that validation definitely helps refine the types of corrections made. It may be that contact with financial and administrators provides the added information students need to correct properly.

4.3: Impact of Validation - Applicants on (not on) the Recipient File and Their Current Eligibility Status

One could hypothesize that validation would have significant impact on the rate at which applicants appear on the recipient file. The following table (Table 4.3) addresses this issue; however, it provides no evidence that validation either facilitates or impedes an applicants process through the system. The table indicates that as of May 1980, approximately 57 percent of all validation applicants appeared on the recipient file. Over 99 percent of these applicants were considered currently eligible for an award; less than one-half of one percent were considered to have received payment erroneously because they are now classified as rejected or uneligible.

In contrast, almost 49 percent of all nonvalidation applicants were on the recipient file as of May, 1980. Again, over 99 percent of these applicants and currently eligible, and only slightly over two-tenths of a percent were paid in error.

It is difficult to use the data in Table 4.3, to ascertain how validation assists or hinders applicants in negotiating the processing system through to payment. Although a higher percentage of validation applicants are on the recipient file than nonvalidation applicants, this does not necessarily mean that validation facilitates payment. One can

TABLE 4.2: POST-SELECTIONS CORRECTIONS,[1] AVERAGE SEI CHANGE AND PERCENT OF APPLICANTS CORRECTING WITH NO SEI CHANGE - VALIDATION AND NONVALIDATION APPLICANTS

Applicant Group	% Correcting Post-Selection/ Eligibility	Average SEI Change	% Correcting With No SEI Change
Validation			
Meeting PEC	38%	128	33.8%
Random	29	37	39.4
Nonvalidation			
Meeting PEC	9	24	64.1
Not Meeting PEC	7	29	59.7

[1]/Corrections refer to corrections made to any field.

conclude from this data, however, that validation applicants seldom stay currently ineligible or rejected while over 55 percent of the nonvalidation applicants are current ineligible or rejected.

In summary, validation is having a significant impact on most PEC subgroups studied, by causing more applicants to correct than overwise would have corrected. In addition, most validation applicants have corrections which result in greater SEI change than nonvalidation applicants. Finally, validation seems to help students correct more selectively (that is, make corrections that affect their SEI, not random corrections that do not affect their potential award), but yet does not significantly impede their progress through the processing and award system.

TABLE 4.3: SUMMARY OF EXPECTED PAYMENT AND CURRENT ELIGIBILITY STATUS OF VALIDATION AND NONVALIDATION APPLICANTS

	Expected to be Paid				Not on Recipient File			
	Total	% Currently Eligible	% Currently Ineligible	% Currently Rejected	Total	% Currently Eligible	% Currently Ineligible	% Currently Rejected
Validation Applicants								
Meeting PEC	56.17%	99.43	.08	0.47	43.82%	94.22	3.80	1.97
Random	58.17	99.61	.05	0.32	41.82	96.51	2.09	1.38
TOTAL	56.60	99.47	.08	0.44	43.39	94.70	3.44	1.85
Nonvalidation Applicants								
Meeting PEC	64.57	99.22	.08	0.68	35.42	97.98	0.63	1.37
Not Meeting PEC	46.09	99.86	.02	0.11	53.90	37.46	44.45	18.07
TOTAL	48.63	99.74	.03	0.21	51.40	43.19	40.30	16.49

4.4: The Interaction of the Impact of Validation and The Processing System Edits

In the analyses thus far, an attempt has been made to separate the impact of validation and the impact of the processing system edits. In Chapter 2, the analysis centered on the immediate response of all applicants subsequent to receiving a processing system edit. In examining the impact of validation, the analysis has focused on the post-selection corrections behavior of validation applicants. The validation process and the edits, however, do not affect two separate populations. About 84 percent of all validation applicants received a processing system edit. Therefore, for a substantial portion of the validation population, the edits are working in tandem with the validation process to elicit complete and accurate data.

Table 2.2, in Chapter 2, showed that the majority of validation applicants' pre-selection corrections were in response to the processing system edits and that these corrections tended to be to the applicants' disadvantage. In particular, PEC applicants made frequent corrections to their disadvantage in response to the edits. In sum, Table 2.2 suggests that the edits had a greater impact than the validation process on the corrections behavior of validation applicants.

In order to establish more fully the impact of the edits relative to the impact of validation, the pre-selection and post-selection corrections behavior of PEC Group A validation applicants was compared. It should be noted that all PEC Group A applicants receive a rejection edit prior to selection. The rejection reason is identical to the selection criterion. In Table 4.4 which follows, each PEC A sub-criterion has been matched with its corresponding rejection reason in order to compare the pre-selection influence of the rejection edit with the post-selection influence of institution validation.

Table 4.4 clearly shows that, regardless of the rejection/selection reason, the edits are having a much more significant impact than validation on validation applicants' corrections. For example, over 60 percent of the validation applicants who received the Social Security

match edits corrected NTI prior to selection, whereas, once selected, less than 2 percent changed NTI. The data from this table strongly suggest that the edits are soliciting the proper corrections prior to selection.

Table 4.5 demonstrates further the interaction of effects of validation and edits. In this table, only the 81,702 validation applicants who were on the recipient file as of May, 1980 and who received an edit prior to selection to one or more of the seven critical fields which financial aid administrators are required to validate are examined. Values for these validation fields were compared at various points in time to determine whether the edits or validation were responsible for the valid data item. The following four categories were used to make this determination:

A -- Applicants whose value at payment does not equal value at selection and does not equal value at edit.

In this type of case, applicants did not make the necessary changes in response to the edit; therefore, the edit was either not effective or partially effective and validation was necessary.

B -- Applicants whose value at payment does not equal the value at selection but equals the value at edit.

In this type of case, the edit caused the applicant to make unnecessary or invalid corrections and validation aided the applicant in reporting accurate data (i.e., the value was changed to what it was at the receipt of the edit).

C -- Applicants whose value at payment is equal to the value at selection and not equal to the value at edit.

In this case, the edit caused the applicant to make the necessary correction before the applicant's selection for validation, rendering validation unnecessary.

D -- Applicants whose value at payment is equal to the value at selection and equal to the value at edit.

Both the edit and validation in this case were unnecessary.

To facilitate the interpretation of this table, the following summarizes the four categories:

A - validation necessary, edits ineffective

B - validation necessary to counteract errors caused by edits

TABLE 4.4: COMPARISON OF CORRECTIONS BEHAVIOR OF VALIDATION APPLICANTS IN RESPONSE TO REJECTION EDITS AND THEIR CORRESPONDING CRITERIA

Rejection Reason/ Selection Criterion	Pre-Selection Corrections In Response To Edit			Unsolicited Post-Selection Corrections		
	Reject Edit	Percent[1/] Correcting	Average Potential[2/] Payment Change	Criteria	Percent[3/] Correcting	Average Potential[4/] Payment Change
Portions Greater Than 120% or 170% AGI	D	34.7%	-$ 70	A-2	11.8%	-$155
Unusual Expenses Greater Than Income or Greater Than $5,000	E	13.3	- 217	A-3	0.5	- 130
Zero, Negative or Less Than $51 Income	F	12.5	- 55	A-4	2.8	- 75
Reported Tax Greater Than Computed Tax by $300 and 50% of Computed Tax	G	38.0	- 24	A-6	19.8	- 221
Unreimbursed Tuition Greater Than 35% of Income	L	27.1	- 222	A-5	2.0	- 35
Applicant's Resources Decreased by $300	R	8.1	- 5	A-9	0.1	0
SSA Match - Blank or Zeroes	S	65.3	- 130	A-10, A-11	1.8	- 198
SSA Match - SS Benefits Reported	T	61.0	- 96	A-10, A-11	1.8	- 198
VA Match - Blank or Zeroes	U	13.8	- 59	A-12, A-13	0.7	- 144
VA Reported and Less Than $131	V	25.7	- 61	A-14, A-15	2.5	- 75
Marital Status, Household Size Blank	7	86.8	+ 155	A-1	3.7	- 93
AGI Blank and Portion Reported	16	60.0	+ 5	A-8	9.2	- 110
Taxes Paid Blank and AGI Greater than Zero	17	51.4	- 10	A-7	7.9	- 48

1/Percentage correcting relevant field on transaction just subsequent to receiving edit.

2/Calculated from receipt of edit to first subsequent transaction.

3/Percentage correcting relevant field between selection transaction and current transaction.

4/Calculated from selection transaction to current transaction.

C - edits effective, validation unnecessary

D - neither edits nor validation necessary

As table 4.5 indicates, validation applicants misreported AGI most often. For 29 percent of the applicants, neither the edits nor validation were necessary to elicit a valid AGI figure. By comparison, the edits and validation were unnecessary in nearly 75 percent of the cases to insure valid PHE.

The edits, regardless of the field, were more responsible for valid data items than the validation process. For over 50 percent of the applicants, the edits were responsible for valid AGI by soliciting large corrections to the applicants' disadvantage. This corroborates the findings from Chapter 2 which suggested that the numerous and restrictive edits which address AGI were working effectively. In certain cases, the edits caused applicants to make erroneous corrections. Three percent of the applicants, for example, improperly changed HS to their advantage in response to the edits.

In over 80 percent of the cases, the items to be validated are accurate before validation takes place. This is due either to the original validity of the fields or the effectiveness of the edits. Validation was needed only 4.4 percent of the time to solicit valid PHE. It was needed most often in regard to AGI, with 17.7 percent of the applicants having an inaccurate AGI at the time of validation. In several cases in Group A where validation was ultimately necessary to elicit a valid field, the edits were partially effective; that is, applicants made corrections in the proper direction in response to the edits.

G-170

QUALITY CONTROL ANALYSIS OF SELECTED ASPECTS OF PROGRAMS ADMINISTERED BY THE BUREAU OF STUDENT FINANCIAL ASSISTANCE

This report was prepared pursuant to Contract No. 300-79-0742. The names of the persons employed by the contractor with managerial or professional responsibility for such work are as follows:

Ms. JoAnn Kuchak, Project Director

Ms. Susan Gray, Senior Project Analyst
(Data Analysis and Report Preparation)

Mr. John Neely, Project Analyst
(Data Analysis and Report Preparation)

Dr. Pedro Saavedra, Technical Manager
(Data Analysis and Report Preparation)

Applied Management Sciences
962 Wayne Avenue, Suite 701
Silver Spring, Maryland 20910

Office of Student Financial Assistance
Department of Education

TABLE 4.5: INTERACTION OF EFFECTS OF VALIDATION AND EDITS ON VALIDATION APPLICANTS

Validation Fields:	Number of Validation Applicants Receiving Edit to This Field	A: Value At Payment Does Not Equal Value At Selection And Does Not Equal Value At Edit			B: Value At Payment Does Not Equal Value At Selection And Equals Value At Edit		
		%	Average Difference in Value From Edit to Selection	Average Difference in Value From Selection to Payment	%	Average Difference in Value From Edit to Selection	Average Difference in Value From Selection to Payment
Adjusted Gross Income	70,310	16.6%	+$3,027	+$4,199	0.8%	+$2,194	-$2,194
Portions Earned	58,140	9.8	-$ 876	+$ 552	0.4	-$ 614	+$ 615
Taxes Paid	66,342	13.5	-$ 356	-$ 653	0.5	+$ 406	-$ 406
Nontaxable Income	46,840	11.1	+$ 756	+$ 911	1.4	+$4,664	-$4,664
Veteran's Educational Benefits	7,638	11.6	-$ 57	-$ 73	1.5	-$1,213	+$1,213
Household Size	5,712	7.0	+ .05	- 1	3.0	+ 2.4	- 2.4
Post-High Enrollment	10,878	2.8	+ .9	+ .05	1.6	+ 1.5	- 1.5

Validation Fields:	Number of Validation Applicants Receiving Edit to This Field	C: Value At Payment Equals Value At Selection And Does Not Equal Value At Edit			D: Value At Payment Equals Value At Selection And Equals Value At Edit		
		%	Average Difference in Value From Edit to Selection	Average Difference in Value From Selection to Payment	%	Average Difference in Value From Edit to Selection	Average Difference in Value From Selection to Payment
Adjusted Gross Income	70,310	53.9%	+$6,316	$0	28.7%	$0	$0
Portions Earned	58,140	38.1	-$1,712	$0	51.7	$0	$0
Taxes Paid	66,342	37.9	-$1,179	$0	47.4	$0	$0
Nontaxable Income	46,840	31.2	-$2,687	$0	56.3	$0	$0
Veteran's Educational Benefits	7,638	35.1	-$ 676	$0	51.8	$0	$0
Household Size	5,712	49.8	+ 0.7	0	40.1	0	0
Post-High Enrollment	10,878	21.2	+ 0.7	0	74.4	0	0

5

ERROR-PRONE MODELLING STUDY

5.1 Purpose and Scope

OSFA is interested in the identification of applicants who are likely to be misreporting on their BEOG applications. This study focuses on the development of an error-prone model (EPM) related to the validation system.

Any investigation of the characteristics of applicants whose forms are likely to contain errors must provide an operational definition of errors which include a means of detecting them. This report addresses errors that can be detected through the validation process employed by institutions in the 1979-1980 academic year.

An overview of this study is found in Chapter 5, along with a description of the process through which the model presented in this report was arrived at. Chapter 6 presents the model in its final form, while Chapter 7 describes each of the groups into which applicants were classified as a result of the analysis. Chapter 8 presents a summary of the EPM study and compares the error-prone model developed here with the existing PEC.

5.2 Key Study Objectives and Questions

Whereas the previous chapter addressed the impact of the existing methods for controlling errors (the edits and validation by pre-established criteria) the broad objectives of the error-prone modeling process undertaken in this investigation are:

- to investigate the characteristics of students most likely to misreport information on their application, and

- to provide a means of identifying such applicants and of minimizing the impact of their misreporting

The specific objectives of the error-prone modeling component include the following:

- to provide a means of estimating the likelihood that an applicant is misreporting
- to provide a method of selecting applicants for validation
- to provide a model that is sensitive to future changes in the BEOG program, (such as different restrictions in eligibility determination)
- to provide information for management improvements specifically related to:
  - -- the validation system
  - -- the edit system
  - -- the original application forms
  - -- reduction of drop-outs among students entitled to funds, and
  - -- a global quality assurance strategy

In more informal terms, the model presented in this report attempts to address the comment, "different corrective measures are needed for different kinds of applicants," by providing a useful operational definition of the term "kinds of applicants."

5.3 Study Population and Samples

The study population for this report is all regular Basic Grant applicants who had an eligible transaction processed by the end of August 1979. This date was chosen after examination of the 1978-1979 data as one assuring that most applicants re-entering the system after selection by that date had been processed at the time the study was conducted. Thus the following sets of applicants were excluded from the study:

- Applicants whose first eligible transaction was processed after August 31, 1979
- Applicants who were never eligible
- Applicants who filed supplemental forms

The samples used in this study are described in Chapter 7, but two issues concerning the population and its relationship to the sample must be discussed here. First of all, the samples through which the

model was developed consisted entirely of applicants selected for validation through the "random" process. This process was not entirely random, since early filers with multiple eligible transactions had a greater chance of being selected. Whatever biases are represented in the random validation sample will be reflected in this study.

The second issue refers to independent applicants affected by the change in the eligibility index formula which became effective May 9, 1979. Unlike the study presented in Part I of this report, Part II retained those applicants. Changes were made to operational definitions to account for the systems-generated transaction these applicants received, but two factors must be taken into account.

First, independents who filed prior to May 9 and were ineligible under the old formula, but became eligible under the new, could not have been selected for random validation if they only filed one transaction. Therefore, although they should have been in the sample, they were not.

Second, the degree to which these applicants' behavior was affected by the different SEI formula and the subsequent change is unknown. For example, the extent to which the change in the SEI computation affected applicants willingness to re-enter the processing system, proceed with validation or change certain fields, is not known. Although these applicants are included in the sample, possible biases due to their presence are taken into account in interpretation of results.

Four samples, drawn from the population of 1979-1980 eligible applicants, were used in this investigation. Each of the first three samples consisted initially of 20,000 applicants, approximately 14,400 dependents and 5,600 independents. Only 10,000 dependents were used in each of the statistical procedures, but cases were weighted accordingly. Each of the four samples used in the analysis is briefly described below.

The working sample consisted of randomly selected validation applicants through the random process whose selection transaction had a process date prior to August 31, 1979. The cutoff date was necessary, since only the first update of the recipient file was available at the time the analysis was conducted. It also enabled us to distinguish as well as possible between applicants who failed to re-enter the system and applicants whose payment transaction was simply in process. An examination of a subsample of the 1978-1979 data revealed that of random validation recipients with no post-selection transaction, 73.8% of those selected by August 31 appeared in the first update of the recipient file, while only 20% of those selected after that date appeared in the first update. Most of the statistical analyses were conducted using this sample.

The replication sample was selected in exactly the same way as the working sample. The majority of the random validation applicants were selected for either the working or the replication sample, but there was no overlap between the two. This sample was used to verify that results obtained from the working sample were not due to chance.

The nonvalidation sample was drawn by simulating the random validation process. The sampling procedure selected 20,000 eligible transactions processed by August 31, 1979 belonging to regular applicants. Once again, the number of dependent applicants was reduced to 10,000. This procedure resulted in not merely a sample of applicants, also a simulated selection transaction for each applicant. The simulation program was executed without knowledge of the systems-generated transaction produced to inform independent applicants of the change in the SEI formula. As a result, some systems-generated transactions were chosen as the simulated selection transaction. Those cases (approximately 200) were eliminated from the sample, resulting in a possible minor underrepresentation of certain kinds of independent applicants in the sample. This sample was used to distinguish the effects of validation from those of edits or spontaneous changes.

The PEC sample consisted of 15,000 applicants selected for validation on account of having met a PEC on a transaction processed prior to August 31, 1979. This sample was used to compare the effectiveness of the PEC to that of the error-prone model.

5.4 Research Methodology

A brief description of the research methodology will be presented here. For a more detailed description see Appendices D and E.

An analytical method known as sequential search analysis or automatic interaction detection (AID) was used to classify applicants into groups which differ as much as possible in a dependent variable relating to their response to validation. The groups are defined in terms of a set of predictor variables. Only variables available at the time of selection were used as predictor variables, including SER fields, variables related to corrections history, and variables obtained through he algebraic manipulation of two or more SER fields. AID first splits the sample into two groups which are as different as possible, and continues this process for each resulting subgroup.

The computer program used in the analyses was THAID, a version of AID designed for nominal scale dependent variables. THAID analyses were conducted using the working sample. The effects of each split were cross-validated through the replication sample. Those splits that could not be replicated were not used in defining the resultng groups. Finally, applicants who met the definitions of the groups, but were not selected for validation were examined to ascertain the reasons for failure to re-enter the system.

5.5 Dependent Variable

Several dependent variables were considered for this investigation. Quantitative variables (selection to payment discrepancies in SEI, scheduled award or expected disbursement, or their absolute values) ignored students who did not re-enter the system after selection for validation. For this reason a decision was made to use a nominal scale dependent variable and a program such as THAID which could handle such variables.

Prior to defining the dependent variable one must define the criterion transaction. This is the transaction which will be assumed to be the most accurate and against which data on other transactions will be compared to estimate the effects of validation, or at least of selection for validation. The criterion transaction for this study is defined as follows:

- If the applicant was paid at or after selection, the latest payment transaction is considered the criterion transaction.
- If the applicant was not paid at or after selection and the latest transaction is different from the selection transaction, then the latest transaction is the criterion.
- If the applicant was not paid on the selection transaction and filed no post-selection transaction, then the criterion is missing.
- Otherwise (i.e., the applicant was not paid at selection and the latest transaction is rejected) the applicant was excluded from the study.

The previously mentioned systems-generated transaction, which notified a small proportion of the applicants of the change in the SEI formula, was ignored in all calculations. If an applicant was paid on this transaction, the payment was assigned for statistical purposes to the latest pre-May 9 transaction.

Given this definition of the criterion transaction, the dependent variable can be defined. This variable will be referred to as TYPE and the term "Type of applicant" will be operationally defined as the value of the variable TYPE for the given applicant. The four types of applicants are:

Type 1: Exact reporters. Applicants for whom the SEI of the criterion transaction is less than 50 points above or below the SEI of the selection transaction.

Type 2: Over-claimers. Applicants for whom the SEI of the criterion transaction is at least 50 points higher than the SEI of the selection transaction.

Type 3: Under-Claimers. Applicants for whom the SEI of the criterion transaction is at least 50 points lower than the SEI of the selection transaction.

Type 4: Missing. Applicants with a missing criterion transaction.

Exhibit 5.1 presents the definition of applicant types in flow-chart form.

The use of expected disbursement, scheduled award or some other standard such as "made corrections to some critical field" were considered as possible alternatives to a nominal scale dependent variable. The first two alternatives were rejected, since institution-related information could have been unavailable if the criterion transaction were not a payment one. The use of critical field corrections was rejected since it would not discriminate between significant and insignificant changes or between over-claimers and under-claimers. The present definition uses information available for all applicants and allows distinctions to be made between kinds of serious misreporters.

5.6 Predictor Variables

Sixty-six predictor variables that were used in the analysis are presented in Exhibit 5.2 with asterisks marking the ones that contribute to defining the model in its final form. Note that some splits could have been defined by more than one variable (e.g., Father's portion is blank or zero could be defined by variables 15 or 17). The variables are in no particular order.

5.7 Strengths and Limitations

One of the major limitations of this approach to error-prone modeling is that one can never be certain of having produced the best possible solution. A different investigator, using different samples, different variables, or forcing a different first split could well come up with a totally different model. What could have been an effective split at the second iteration may never appear in light of a different split at the first.

THAID has a tendency to select splits that come close to dividing a sample or subsample in an uneven division. For this reason criteria which define only a very small percentage of the population would be unlikely to appear in the model. The one tenth of one percent of the applicants defined by a given PEC could be expected to be scattered among the thirty-seven groups resulting from the study. For this reason the model should not be used to the exclusion of other criteria that have also proven effective.

EXHIBIT 5.1: DEFINITION OF DEPENDENT VARIABLE

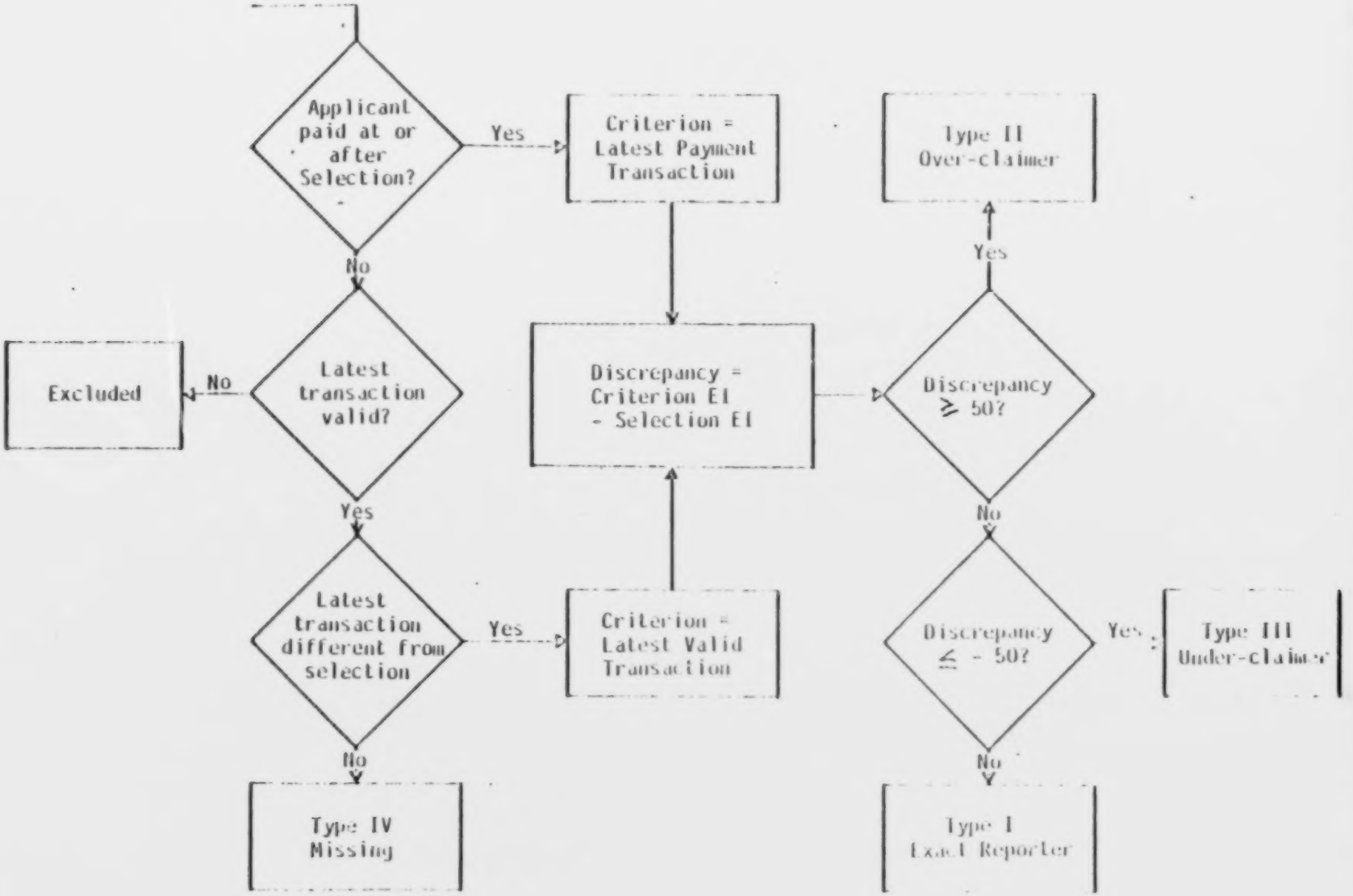


EXHIBIT 5.2: PREDICTOR VARIABLES

1. Citizenship (citizen, resident or blank)
2. Bachelor's degree (no or blank)
*3. Tax filing status (based on tax form, estimated, did not have to file or blank)
4. Number of prior transactions
*5. M.D.E. source (SACC was coded the same as CSS)
6. Marital status
7. Discrepancy between household size and exemptions
8. Household size
*9. Exemptions
10. Marital status
11. Social Security Benefits
12. Non-taxable other than Social Security
13. Non-taxable income
*14. Adjusted Gross Income
15. Father's portion
16. Mother's portion
*17. Presence or absence of both sources of earned income (both absent, father's only, mother's only, both present)
*18. Taxes paid
*19. Taxes as a proportion of AGI (blanks converted to zeros)
20. Post-high school education (members of household in)
21. Unreimbursed tuition
22. Medical of dental expenses
23. Casualty-theft losses
24. Itemized deductions
*25. Savings
26. Applicant's resources
27. Veteran's benefits (amount only)
*28. House value
29. House debt
30. Investment value
*31. Date first application signed
*32. Data selection transaction processed
33. Date of birth
*34. Type of school (first choice)
*35. Control of school (first choice)
36. Presence or absence of second choice school
37. Congruence between state listings (whether state of legal residence, state in applicant's address and state of first choice school coincide; one value for each possibility including missing school)
38. Number of blanks and zeros in fields calling for dollar responses
*39. Eligibility index
*40. Total income (N.T.I. + A.G.I. + amount received from veteran's benefits in a year after converting blanks to zeros)

EXHIBIT 5.2: PREDICTOR VARIABLES (CONTD.)

*41. Year in school
42. Number of assumption comments
43. Unusual expenses (Medical or dental + casualty-theft losses)
*44. Assets (House, farm, business and investment values minus debts, plus savings and applicant's resources)
45. Father's portion divided by total income
46. Total income divided by (total income + assets)
47. Applicant's resources divided by parent's total income
48. Total debts divided by (house, investment, farm and business values + savings)
*49. Non-taxable income divided by total income
50. Number of critical fields having received rejection comments in the past
51. Number of changes made to critical fields
52. Number of verifications in selection transaction
53. Number of fields producing assumption comments across transactions
54. SEI difference between immeidate prior transaction and present transaction (SEI was computed if previous transaction was not valid)
55. Previous ineligible transaction?
56. Ever changed marital status?
57. Ever changed household size?
58. Ever changed U.S. tax figures?
59. Ever changed exemption?
60. Ever changed AGI?
61. Ever changed taxes paid?
62. Ever changed model?
63. Ever a change in scheduled award?
*64. Absolute value of the difference between highest and lowest SEI achieved (included computed for rejections)
65. Number ever missing among the following fields: bachelor's degree, household size, U.S. tax figures, exemptions, A.G.I., taxes paid, date signed, year in school
*66. Model (incorporated by forced division of sample, but not chose by THAID).

While other methods such as discriminant analysis may identify small effects defined by a single variable, they are likely to miss larger effects defined by a combination of two or more variables. THAID is particularly effective at identifying predictors which interact to produce an effect on the dependent variable. The sequential search method also yields results which are more easily interpreted and which lend themselves to a greater variety of remedial actions.

6

THE ERROR-PRONE MODEL

6.1 Results of the Validation and Nonvalidation Sample Comparisons

This section presents the overall findings of the analysis with respect to the validation and nonvalidation groups. The general findings corroborate those reported in Part I of this report, namely that selection for validation is accompanied by different corrections behavior from the rest of the applicant population.

Among random validation applicants included in the working sample, 8.9 percent were over-claimers (i.e. made corrections to their SER resulting in at least a 50 point increase in SEI from selection to criterion transaction), 4.5 percent were under-claimers and 35.6 percent failed to re-enter the system (these figures were calculated weighted dependent sample). The corresponding figures for the replication samples were very similar, at 8.7 percent, 4.4 percent, and 36.1 percent, respectively.

For the nonvalidation sample, however, only 1.3 percent corrected to their disadvantage and 0.7 percent corrected to their advantage. The percentage failing to re-enter the system was 31.2 percent. In other words 20.8 percent of applicants who re-enter the system after selection for validation make corrections affecting their SEI by 50 points or more, but only 2.9 percent of those not selected for validation do so.

6.2 Description of Model Results

This section presents a global description of the error-prone model, nighlignting its most important features and discussing the major findings associated with the model.

Thirty-seven groups, 25 dependent and 12 independent were identified through the sequential search procedure. Twenty variables suffice as definers of the 37 groups:

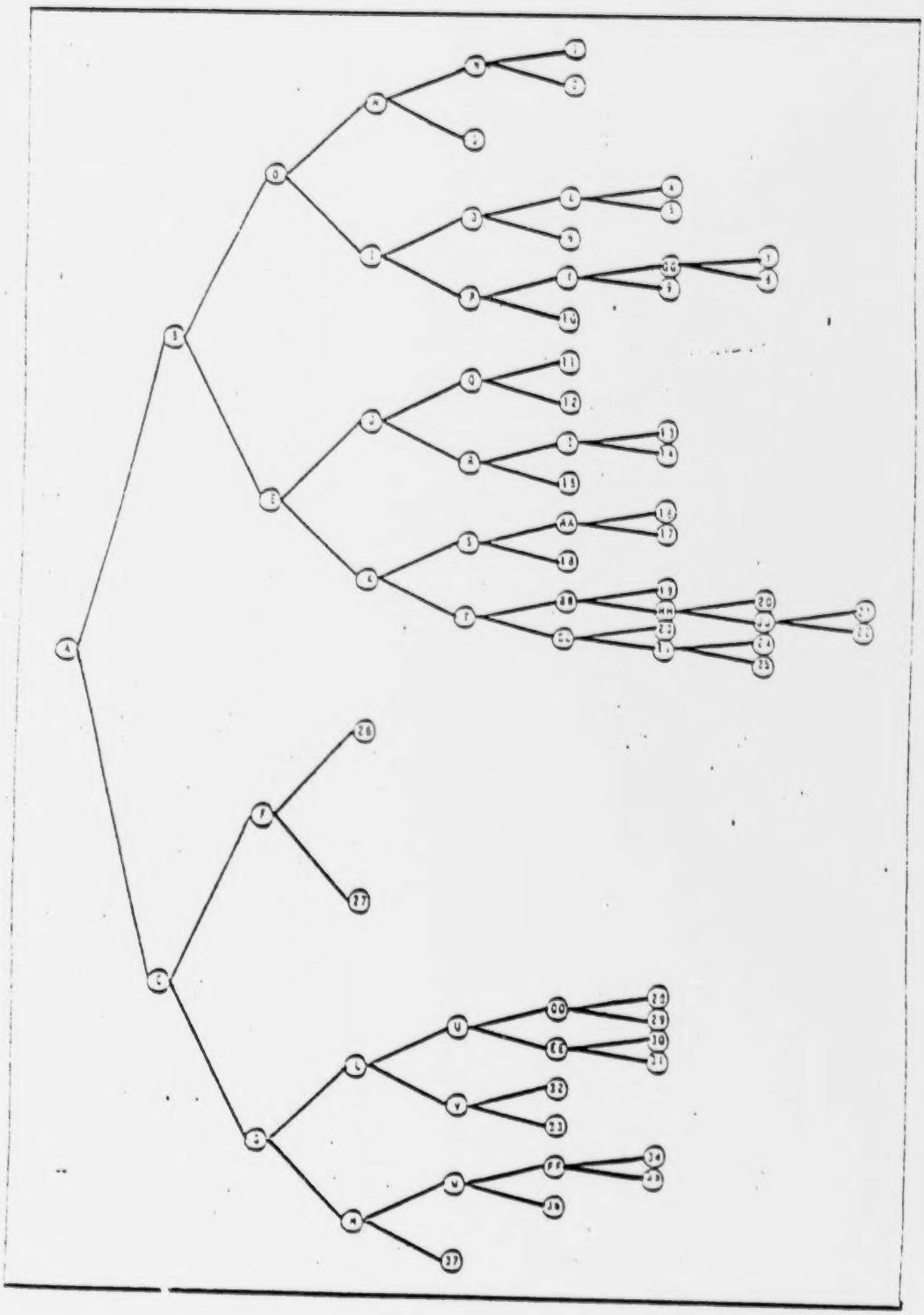
- one relates to corrections behavior (SEI change),
- five are algebraic manipulations of SER fields (SEI, total income, assets, taxes as a proportion of AGI, and NTI as a proportion of total income),
- one is a logical statement relating two fields (presence or absence of each portion earned, with four possibilities),
- two are dates (date application signed and date transaction processed),
- two are related to school choice (type and control),
- one is related to application form (MDE source),
- One is answered indirectly by the applicant (model or dependency status), and
- seven are answered directly in the application (AGI, savings, house value, tax filing status, taxes paid, exemptions and year in school).

Exhibits 6.1 and 6.2 provide two different ways of viewing the model. Exhibit 6.1 presents the tree indicating the splits that led to the 37 groups. Letters represent nodes (subsamples which were subsequently split) and numbers represent the 37 groups. As can be seen, eight splits were the most that were necessary to create the most complex groups, while three sufficed for the simplest.

Exhibit 6.2 simply lists the defining characteristics of each group.

Before considering the misreporting patterns of the different groups, the reader should be made aware of certain trends in the splits producing the model. Applicants whose MDE source was BEOG had similar patterns to those whose MDE source was PHEAA, while ACT applicants resembled CSS applicants (SACC applicants were coded as if they were CSS). ACT and CSS applicants produced more splits (and therefore more groups) than did BEOG and PHEAA applicants. Twenty groups were composed entirely of ACT and CSS applicants, seven of BEOG and PHEAA applicants, and ten had applicants from all four sources.

EXHIBIT 6.1: ERROR-PRONE MODEL TREE DIAGRAM



6.3

EXHIBIT 6.1 (CONT.): NODES IN TREE DIAGRAM

Node:	Definer of Latest Split
A	All applicants
B	Dependents
C	Independents
D	Estimated taxes
E	Used tax form, did not have to file or left space blank
F	BEOG or PHEAA
G	ACT or CSS
H	SEI not over 400
I	SEI over 400
J	BEOG or PHEAA
K	ACT or CSS
L	SEI = 0
M	SEI greater than 0
N	Savings 0 or blank
O	Taxes paid not over $2,000
P	Taxes paid over $2,000
Q	House value = 0 or missing
R	House value greater than 0
S	SEI not over 200
T	SEI over 200
U	Exemptions = blank, 0 or 1
V	Exemptions = greater than 1
W	Processed by April 30
X	SEI never has changed
Y	Both parents' positions are greater than zero
Z	First choice school is university or blank
AA	Total income less than or equal to $10,000
BB	Taxes less than 5% of AGI or both figures zero
CC	Taxes over 5% of AGI
DD	Processed by April 30
EE	Processed after April 30
FF	Taxes paid not over $500
GG	SEI less than or equal to 1200
HH	Taxes paid greater than zero
II	Taxes over 15% of AGI
JJ	SEI not over 600

EXHIBIT 6.1 (CONT.): GROUPS IN TREE DIAGRAM

Group:	Definer of Latest Split
1.	Father's portion = 0 or blank
2.	Father's portion greater than 0
3.	Savings greater than 0
4.	Processed by May 31
5.	Processed after May 31
6.	SEI changed at least once
7.	First choice school is university
8.	First choice school is blank or not university
9.	SEI greater than 1200
10.	At least one parent's portion is zero or blank
11.	First choice school control is blank or public
12.	First choice school is private or proprietary
13.	BEOG
14.	PHEAA
15.	First choice school listed, but not a university
16.	NTI less than or equal to 50% of total income (or both zero or one negative)
17.	NTI over 50% of total income
18.	Total income over $10,000
19.	Taxes paid = blank or 0
20.	SEI not over 600
21.	Assets not over $30,000
22.	Assets over $30,000
23.	Taxes over 15% of AGI
24.	AGI not over $25,000
25.	AGI over $25,000
26.	SEI = 0
27.	SEI greater than 0
28.	AGI not over $2,000
29.	AGI over $2,000
30.	Processed by July 31
32.	Year in school = blank or 1
33.	Year in school = 2,3,4 or 5
34.	Signed by February 28 or date missing
35.	Signed after February 28
36.	Taxes paid over $500
37.	Processed after April 30

EXHIBIT 6.2: DEFINITION OF THE 37 GROUPS

Groups 1-10: Dependents who estimated taxes

1. SEI not over 400, no savings or father's portion.
2. SEI not over 400, no savings, but father's portion greater than 0.
3. SEI not over 400, savings greater than 0.
4. SEI over 400, taxes not over $2,000, SEI has not changed, processed by May 31.
5. SEI over 400, taxes not over $2,000, SEI has not changed, processed after May 31.
6. SEI over 400, taxes not over $2,000, SEI has changed.
7. SEI 401-1200, taxes over $2,000, both portions greater than 0, first choice is university.
8. SEI 401-1200, taxes over $2,000, both portions greater than 0, first choice missing or not university.
9. SEI over 1200, taxes over $2,000, both portions greater than 0.
10. SEI over 400, taxes over $2,000, one or both portions is 0 or blank.

Groups 11-15: Dependents, did not estimate, MDE is BEOG or PHEAA

11. House value 0 or blank, first choice school public or blank.
12. House value 0 or blank, first choice school private or proprietary.
13. MDE is BEOG, house value greater than 0, university or no first choice.
14. MDE is PHEAA, house value greater than 0, university or no first choice.
15. House value greater than 0, first choice school is listed, but not a university.

Groups 16-25: Dependents, did not estimate, MDE is ACT or CSS

16. SEI not over 200, total income not over $10,000, NTI is not over 50% of total income.
17. SEI not over 200, total income not over $10,000, NTI over 50% of total income.
18. SEI not over 200, total income over $10,000.
19. SEI over 200, taxes paid 0 or blank.
20. SEI 201-600, taxes greater than 0, but not over 5% of AGI.
21. SEI over 600, taxes greater than 0, but not over 5% of AGI, assets not over $30,000.
22. SEI over 600, taxes greater than 0, but not over 5% of AGI, assets over $30,000.
23. SEI over 200, taxes 5 to 15% of AGI.
24. SEI over 200, taxes over 15% of AGI, AGI not over $25,000.
25. SEI over 200, taxes over 15% of AGI, AGI over $25,000.

EXHIBIT 6.2: DEFINITION OF THE 37 GROUPS (CONT.)

Groups 26-27: Independents, MDE is BEOG or PHEAA

26. SEI = 0.
27. SEI greater than 0.

Groups 28-37: Independents, MDE is ACT or CSS

28. SEI = 0, exemptions are 1 or blank, processed by April 30, AGI not over $2,000.
29. SEI = 0, exemptions are 1 or blank, processed by April 30, AGI over $2,000.
30. SEI = 0, exemptions are 1 or blank, processed May 1 to July 31.
31. SEI = 0, exemptions are 1 or blank, processed after July 31.
32. SEI = 0, exemptions greater than 1, year in school is 1 or blank.
33. SEI = 0, exemptions greater than 1, year in school greater than 1.
34. SEI greater than 0, processed by April 30, taxes not over $500, signed by February 28 or date missing
35. SEI greater than 0, processed by April 30, taxes not over $500, signed after February 28.
36. SEI greater than 0, processed by April 30, taxes over $500.
37. SEI greater than 0, processed after April 30.

Type and control of institution contributed to the definition of all five dependent BEOG and PHEAA groups, and of two of the groups including all four sources. Type and control of institution did not seem to make a difference for ACT and CSS applicants, since they entered into the definition of none of the twenty groups composed entirely of applicants from these two sources.

Aside from SEI (which split the independents), information related to tax forms was not at all useful in predicting misreporting for BEOG and PHEAA applicants. Dates of processing and signing also failed to predict for this subsample. On the other hand, SEI, dates (for independents) and information that could be verified directly from the tax forms (AGI, taxes paid, exemptions, etc.) proved to be the best predictors for ACT and CSS applicants. Among applicants who estimated taxes (the ten groups containing applicants from all four sources) EI, portions (presence or absence) taxes paid and savings proved the best predictors.

Table 6.1 shows the distributions on the dependent variable for the two random validation and the nonvalidation samples. The column marked % lists the percentage of the total sample represented by the group. The other four columns represent the percentage of each type in each group.

As can be seen, Group 7 has the highest percentage of overclaimers, but most of the groups made up of dependents who estimated taxes are also relatively high. Some of these (like Groups 2 and 3) have four and five times as many over-claimers as under-claimers, while others (like Groups 4 and 6) have approximately equal numbers of both types.

Group 24, characterized by a high tax rate for the income of its members, is the only group among dependents who did not estimate taxes with a high proportion of over-claimers. Group 21 has the highest proportion of under-claimers.

TABLE 6.1: GROUP SUMMARY CHART

Group	Working Sample %	I	II	III	IV	Replication Sample %	I	II	III	IV	Non Validation Sample %	I	II	III	IV
1	1.5	53.3	8.9	2.8	35.0	1.5	53.8	8.0	4.2	34.0	1.4	67.5	2.0	0.0	30.5
2	1.2	37.3	19.3	3.6	39.8	1.3	44.4	18.9	2.2	34.4	1.1	75.5	2.5	0.6	21.4
3	2.4	51.5	21.2	5.0	22.4	2.5	49.3	21.9	2.9	25.9	2.2	76.6	1.6	0.6	21.1
4	2.1	35.4	24.3	21.5	18.8	1.9	38.5	21.1	16.3	24.1	2.2	67.6	6.0	2.0	24.4
5	1.3	36.0	19.4	17.1	27.4	1.2	40.7	24.0	12.6	22.8	1.1	69.7	0.6	1.9	27.7
6	1.5	46.2	12.0	12.5	29.3	1.9	42.9	16.8	12.7	27.6	1.1	74.2	1.3	1.3	23.3
7	0.7	39.4	35.6	9.6	15.4	0.8	40.0	33.0	6.1	20.9	0.9	71.3	2.5	0.8	25.4
8	0.8	55.8	21.2	4.4	18.6	0.8	44.6	25.9	5.4	24.1	0.6	64.4	4.4	1.1	30.0
9	1.0	49.6	13.7	5.8	30.9	1.3	43.0	15.6	10.6	30.7	1.3	75.4	5.0	1.1	18.4
10	1.5	32.2	26.8	6.8	34.1	1.2	39.0	21.5	7.6	32.0	1.4	64.9	6.9	2.1	26.1
11	3.8	38.0	3.2	1.7	57.1	4.2	35.7	4.8	3.8	55.7	4.4	36.8	0.2	0.2	62.8
12	1.5	51.6	5.6	3.8	39.0	1.6	52.0	3.6	2.7	41.6	1.5	64.5	1.4	0.5	33.6
13	2.5	49.6	8.5	4.0	38.0	2.3	48.4	4.4	4.7	42.5	2.2	58.0	0.3	1.9	39.7
14	1.1	27.3	12.7	4.0	56.0	1.1	30.6	7.5	0.7	61.2	1.2	39.4	0.0	1.2	59.4
15	4.7	55.4	7.6	4.7	32.2	4.8	51.1	8.8	4.7	35.4	4.7	70.3	0.6	0.5	28.6
16	6.2	65.3	8.2	0.6	25.9	5.7	64.1	6.9	1.0	28.0	6.2	77.1	0.9	0.1	21.9
17	4.7	60.8	5.7	0.8	32.7	4.5	60.8	5.0	1.3	33.0	4.6	72.7	0.6	0.0	26.7
18	3.6	51.4	8.4	2.0	38.2	3.9	56.2	9.9	1.5	32.5	3.5	74.8	1.0	0.4	23.8
19	3.3	42.0	12.8	17.5	27.7	3.5	47.0	10.2	13.9	28.8	3.3	76.4	0.9	3.1	19.6
20	1.7	57.9	6.2	5.0	31.0	1.7	57.0	9.3	7.6	26.2	1.9	74.5	0.7	1.1	23.6
21	1.0	40.0	8.1	22.2	29.6	0.9	40.3	12.6	23.5	23.5	0.8	66.7	0.9	1.8	30.7
22	0.9	56.9	6.5	11.4	25.2	0.8	48.1	8.5	9.4	34.0	0.7	73.9	2.2	3.3	20.7
23	20.7	59.4	8.9	5.0	26.7	20.3	58.2	7.9	5.1	28.8	21.5	77.0	1.6	0.8	20.5
24	0.8	33.0	25.9	4.5	36.6	0.7	34.6	29.8	4.8	30.8	0.7	75.2	5.0	1.0	18.8
25	1.3	53.8	14.8	2.7	28.6	1.4	54.4	11.0	2.1	22.5	1.4	75.9	0.5	1.5	22.1
26	8.3	43.9	1.9	0.0	54.2	8.8	43.1	1.8	0.0	55.1	7.6	49.4	0.2	0.0	50.4
27	2.5	28.6	6.9	8.5	56.0	2.6	33.3	8.2	5.6	52.9	2.3	50.0	0.9	0.5	48.5
28	3.2	51.8	1.9	0.0	46.3	2.9	51.0	3.6	0.0	45.4	2.9	49.6	1.2	0.0	49.2
29	1.2	36.7	4.0	0.0	59.3	1.2	33.9	7.1	0.0	59.0	1.1	22.2	2.4	0.0	75.5
30	2.6	59.5	3.1	0.0	37.4	2.5	59.7	4.4	0.0	35.9	2.5	69.2	0.4	0.0	30.3
31	0.6	46.0	1.6	0.0	52.4	0.8	50.6	2.4	0.0	47.1	0.9	63.9	0.5	0.0	35.5
32	2.3	56.6	3.5	0.0	39.9	2.3	54.8	2.6	0.0	42.6	2.1	65.1	0.5	0.0	34.4
33	3.1	65.2	3.3	0.0	31.5	2.9	56.8	4.2	0.0	29.1	2.7	71.3	1.3	0.0	27.4
34	0.5	21.1	12.3	10.5	56.1	0.6	32.7	15.0	8.8	43.4	0.5	49.0	1.0	2.0	48.0
35	0.3	43.5	7.2	10.1	39.1	0.4	34.2	6.6	6.6	52.6	0.4	45.2	1.4	0.0	53.4
36	0.4	12.7	17.7	2.5	67.1	0.3	16.4	11.5	6.6	65.6	0.3	32.4	1.5	0.0	66.2
37	2.7	45.3	9.5	5.6	39.6	2.7	47.4	8.1	7.5	37.1	2.7	61.3	1.9	0.7	36.1
Total	100.0	51.1	8.9	4.5	35.6	100.0	50.9	8.7	4.4	36.1	100.0	66.7	1.3	0.7	31.2

I = Exact reporters

II = Over-claimers

III = Under-claimers

IV = Missing

Independents had proportionately fewer misreporters, but there were some independent groups with a large proportion missing and a relatively high rate of misreporting among those who did re-enter the system. The missing rate could not be attributed to a deterrent effect of validation, since it was generally replicated in the nonvalidation sample. It is possible that the change in SEI formula (which would have affected many of these applicants) could have led financial aid administrators to recommend that applicants wait until they received the system generated SER. If so, they may be delaying their re-entry. Alternately, some may nave felt their grant was too small (under the old formula) and made other plans.

While no nonvalidation group had more than 7 percent making corrections which raised their SEI (the equivalent of over-claimers for the validation groups) some of the groups with the highest proportions in tnis category were also groups with high percentages of over-claimers. In other words, validation groups with high corrections rates tend to correspond to nonvalidation groups with relatively high proportions of spontaneous corrections, but the rates of corrections among the nonvalidation groups are much lower than those of the corresponding validation groups.

6.3 Applicants Who Did Not Re-enter the System

In order to examine the extent to which applicants may have been intimidated by selection, the percentage of missing applicants for each nonvalidation group was subtracted from the percentage of missing applicants in one of the validation groups (whichever yielded the most conservative results). Table 6.2 presents this information. Four groups yielded results over 10 percent. Two of these (Groups 2 and 24) had high rates of over-claimers. The third group (Group 9) combined estimating taxes with a high SEI, possibly leading its members to decide that validation was not worth the trouble since the award would have been low. The fourth group (Group 31) had selection transactions processed in August, and the delay in getting validation materials together may have made these applicants miss the first update of the recipient file.

One group had significantly fewer missing applicants when selected for validation than otherwise. This group is the one most likely to have been affected by the formula change in the sense that their old EI would have been high, while their new one would be zero. Seventy-five percent of this group's nonvalidation applicants failed to re-enter the system, as opposed to 59 percent of the ones in which validation sample. In this case validation may have been an incentive to remain in the system.

When one compares the figures in the first column with percentages of misreporters for non-missing cases no clear pattern emerges. There does seem to be a greater tendency for validation applicants to be missing, but the reasons are probably varied and difficult to determine.

6.4 Misreporting by Critical Field

As was stated in Part I, the fields most frequently changed after selection were NTI, AGI and taxes paid. Table 6.3 reports the percentage of non-missing applicants whose SEI and selection transaction values in these fields were too high or too low. The average amount by which each of these fields was misreported is given in Table 6.4.

An examination of those groups with high percentages underreporting NTI reveals that they are not groups with large numbers of serious misreporters. Groups 11 and 17 had 20.3 percent and 25.4 percent underreporting their NTI, but their percentage of over-claimers (counting only non-missing applicants) were only 7.5 percent and 9.3 percent respectively, well under the average of 13.8 percent. Only for Group 19 does NTI seem to be a possible source of serious error, and here it affects both over-claimers and under-claimers.

AGI is another matter altogether. Group 8 had 50 percent of its non-missing members misreporting AGI to their advantage (though not one of them was found to have underreported NTI). Group 4 had 71.4 percent of those applicant who re-entered the system making corrections to AGI after selection. Table II-4 shows that the mean discrepancy is over $1,000 for seven groups, and this includes all applicants who re-entered,

TABLE 6.2: EFFECTS OF VALIDATION ON RE-ENTERING AND ON SEI CHANGE

Group	Difference In Percentage Missing* (Validation-Non-validation)	Serious Misreporters as Percentage of Non-missing**	Total Misreporters As Percent of Non-missing***
1	3.5	17.9	25.9
2	13.0	38.0	56.0
3	1.2	33.7	44.7
4	-.03	56.4	80.0
5	-.04	50.4	65.3
6	4.3	34.7	42.9
7	-4.5	53.4	70.5
8	-0.9	31.5	61.9
9	12.3	28.1	46.9
10	5.9	51.1	63.0
11	-5.7	11.5	17.2
12	5.4	15.4	19.2
13	-1.7	20.1	25.1
14	1.8	37.9	47.0
15	3.6	18.2	22.5
16	4.0	11.9	16.1
17	6.0	9.6	14.4
18	8.7	16.9	25.3
19	8.1	41.9	49.7
20	2.6	16.2	25.0
21	1.1	43.2	51.6
22	4.5	23.9	37.0
23	6.2	19.0	27.5
24	12.0	47.9	60.6
25	0.4	24.6	29.3
26	3.8	4.3	4.4
27	4.4	35.1	38.8
28	-2.9	3.5	4.4
29	-16.2	9.9	11.9
30	5.6	5.0	5.6
31	11.6	3.4	5.1
32	5.5	5.8	6.2
33	1.7	4.8	5.3
34	-4.6	52.0	56.0
35	-0.8	28.6	28.6
36	-0.6	61.5	65.4
37	1.0	25.0	25.6

* Most conservative estimate, validation minus nonvalidation.
** SEI change of 50 points or more, working sample.
*** SEI change of 1 or more points, working sample.

TABLE 6.3: PERCENTAGE MISREPORTING MAJOR FIELDS

	NTI		AGI		Taxes		SEI	
Group	Too High	Too Low	Too High	Too Low	Too High	Too Low	Too High	Too Low
1	5.0	15.1	9.4	26.6	8.6	10.1	7.9	18.0
2	1.0	6.0	20.0	38.0	16.0	26.0	11.0	35.0
3	5.3	15.2	12.1	17.7	13.3	15.5	11.7	33.0
4	4.3	5.6	24.4	47.0	27.8	35.9	36.8	43.2
5	4.7	5.5	15.0	35.4	19.7	33.1	29.1	36.2
6	8.8	4.8	7.5	16.3	8.2	13.6	21.8	21.1
7	0.0	1.1	18.2	39.8	22.7	36.4	21.6	48.9
8	2.2	0.0	7.6	50.0	21.7	32.6	13.0	48.9
9	0.0	0.0	7.3	31.3	17.7	25.0	17.7	29.2
10	1.5	3.7	8.9	38.5	27.4	22.2	14.1	48.9
11	6.6	20.3	4.8	6.2	5.3	2.2	5.7	11.5
12	10.0	8.5	2.3	8.5	2.2	6.2	7.7	11.5
13	6.8	10.5	3.7	8.2	5.9	3.7	8.2	16.8
14	0.0	16.7	6.1	7.6	9.1	10.6	12.1	34.8
15	5.0	10.4	2.9	6.1	5.9	2.3	9.0	13.5
16	1.5	12.4	4.5	14.7	6.1	6.4	2.2	13.9
17	8.9	26.4	2.1	6.7	0.2	1.1	2.5	11.9
18	3.6	11.0	4.5	13.0	7.8	7.5	5.5	19.8
19	13.2	18.0	9.0	9.9	0.0	10.8	27.5	22.2
20	2.4	2.4	4.8	5.4	5.4	6.0	12.6	11.4
21	9.5	11.6	7.4	14.7	3.2	22.1	33.7	17.9
22	8.7	3.3	8.5	6.5	7.6	9.8	23.9	13.0
23	1.4	2.6	3.9	10.3	8.1	5.1	9.4	18.1
24	1.4	1.4	5.3	23.9	40.8	5.6	8.5	52.1
25	0.0	0.8	3.8	12.3	12.3	2.3	6.2	23.1
26	4.6	12.0	3.0	8.9	3.8	2.0	0.0	4.4
27	5.0	5.0	6.8	10.4	5.4	9.0	22.1	16.7
28	4.7	7.6	1.5	18.7	5.8	3.2	0.0	4.4
29	0.0	2.0	6.9	20.8	17.8	6.9	0.0	11.9
30	3.1	6.2	2.2	13.0	4.0	4.0	0.0	5.6
31	1.7	6.8	8.5	6.8	5.1	0.0	0.0	3.1
32	5.1	13.9	2.2	8.8	1.8	2.2	0.0	6.2
33	3.1	11.0	3.6	8.4	2.6	2.1	0.0	5.3
34	4.0	10.0	8.0	24.0	8.0	18.0	26.0	30.0
35	11.9	7.1	7.1	2.4	0.0	0.0	16.7	11.9
36	3.8	3.8	7.7	26.9	30.8	11.5	11.5	53.8
37	4.0	5.2	2.2	9.0	7.1	5.2	9.3	16.4

not just those who changed the field. (The average discrepancy for Group 2 was $2,230, but if one counts only applicants who corrected this field, the average correction was $3,845 in either direction.) AGI is therefore one of the fields most frequently changed, and probably the one causing most errors. It is also one of the easiest items for financial aid administrators to validate.

Taxes paid, which is also easy to validate, was a high error rate field. Among Group 24 applicants, 40.8 percent of those not missing overreported their taxes (particularly disturbing given that they claimed to not have estimated). Group 4 had the largest proportion of misreporters in this field (62.7%).

Changes in AGI were not regularly accompanied by changes in portions earned. With the exception of Group 2, 41 percent of which changed father's portion, the percentage changing portions in any group seldom reached 20 percent.

6.5 Impact of Misreporting

Table 6.5 presents a breakdown of EI and scheduled award discrepancies between criterion and selection transactions. As can be seen, for some groups (e.g. Group 5) over-claimers and under-claimers cancel each other out, even though the average award may be off by over $100. For other groups (e.g. Group 10) over-claimers far outweigh under-claimers producing a mean net differences of $76 per applicant. Assuming that expected disbursement is 94 percent of scheduled award (as was the case for the 1978-1979 academic year) and that the proportion of non-missing applicants represented by this group is the same as its representation in the recipient file the amount saved by validating all Group 10 applicants over validating none would be over $2,000,000. This is not a net savings amount, since it does not take into account the fact that some are being validated already, or that a deterrent effect might produce additional savings.

Group by group recommendations and interpretations are presented in Chapter 7 and implications for validation are suggested in Chapter 8.

TABLE 6.4: MEAN NET AND ABSOLUTE DISCREPANCIES IN CRITICAL FIELDS

Group	NTI		AGI		Taxes	
	Net*	Abs.	Net*	Abs.	Net*	Abs.
1	179	614	329	591	13	45
2	87	134	1,230	2,230	103	202
3	325	487	455	1,002	41	128
4	-174	282	421	1,493	166	366
5	-18	184	264	1,143	230	355
6	-448	548	98	608	56	110
7	22	22	802	1,193	68	298
8	-4	4	491	979	-67	288
9	0	0	333	516	73	239
10	131	148	567	1,165	-173	426
11	427	708	242	660	-26	39
12	-8	359	62	373	35	49
13	119	301	95	229	5	49
14	477	477	-279	421	-8	154
15	174	363	-57	353	-57	83
16	350	397	670	855	15	81
17	473	769	256	320	10	11
18	264	345	456	615	-20	142
19	-85	773	100	583	144	144
20	18	87	17	201	17	42
21	-193	633	-92	564	280	296
22	-122	285	-154	250	98	117
23	-11	125	42	253	-8	69
24	-117	239	85	1,145	-312	447
25	22	22	168	2076	-128	134
26	177	263	86	172	-13	22
27	-27	164	49	312	16	36
28	72	174	301	317	7	19
29	13	13	271	370	0	60
30	90	140	390	413	27	40
31	89	124	-61	181	-15	15
32	326	382	172	257	2	17
33	198	300	298	409	15	26
34	-32	187	88	457	62	100
35	22	172	-48	85	0	0
36	16	76	294	306	-83	108
37	-56	154	125	236	4	49

* Net discrepancies are averaged without taking the absolute value, so that overreporters and underreporters could cancel each other out. Selection value is subtracted from criterion value.

TABLE 6.5: MEAN NET AND ABSOLUTE SEI AND AWARD DISCREPANCIES

Group	SEI		Award	
	Net*	Abs.	Net**	Abs.
1	23	43	12	21
2	75	96	28	35
3	76	98	46	53
4	25	181	14	125
5	-15	143	0	107
6	12	127	2	82
7	75	111	68	99
8	49	85	50	78
9	29	50	30	52
10	92	123	76	100
11	22	59	3	20
12	8	38	10	29
13	21	45	12	30
14	27	82	22	58
15	7	51	3	31
16	55	57	26	27
17	29	31	12	13
18	47	55	23	26
19	-8	150	10	89
20	3	29	4	14
21	-62	127	-43	107
22	-27	66	-13	51
23	14	50	10	38
24	55	109	57	95
25	57	71	33	44
26	25	25	9	9
27	-19	169	-16	77
28	18	18	6	6
29	57	57	29	29
30	38	38	26	26
31	13	13	0	0
32	52	52	21	21
33	54	54	21	21
34	97	234	75	133
35	15	141	39	77
36	155	198	62	103
37	155	198	62	103

* Criterion-selection.
** Selection-criterion.

7

DESCRIPTION AND INTERPRETATION OF THE GROUPS

This chapter describes each group in terms of its misreporting patterns and selection transaction profile. An attempt will be made to interpret and identify possible reasons for misreporting and to suggest corrective action. Each description should be read in conjunction with Tables 6.1 through 6.5.

At this point a few terms should be reviewed. The term over-claimer refers to an applicant whose SEI was at least 50 points lower on the selection transaction than on the criterion transaction. Under-claimer refers to an applicant whose SEI was at least 50 points higher at selection than at criterion. Misreporter includes both over-claimers and under-claimers. The net discrepancy on a given field between criterion and selection transaction refers to the average of the figure at criterion minus the figure at selection. Absolute award difference and absolute discrepancy on a given field refers to the average of the absolute value of the figure at criterion minus the figure at selection. Gains and losses may cancel each other out in computing net discrepancy of net award difference, but will not in computing absolute award difference or absolute discrepancy.

Group 1: dependents who estimated taxes, had SEI not over 400, no savings and father's portion mission or 0. This group had the lowest rates of misreporting among the ten groups of dependents who estimated taxes. They are a low income group, where the mother's earned income accounts for most of the AGI. This group had the largest percentage of

corrections to mother's portion (26.6%), but its rate of misreporting identified through validation was below that of the total sample. On the other hand, since this group consists largely of applicants with divorced or separated parents the possibility of an applicant providing the wrong definition of parent (i.e., listing their mothers when they lived a longer period with their fathers) is a real one that validation may not identify. Quality control studies might wish to investigate this issue. This group should not be selected for validation, but should be incorporated in any remedial action pertaining to estimation of taxes.

Group 2: dependents who estimated taxes, had SEI not over 400, no savings and father's portion greater than 0. This group differs in its defining characteristics from the previous one only in that father's portion was present. Its misreporting pattern, however, is very different. Over-claimers constitute 19.3 percent of the applicants in this group, and 32.1 percent of those not missing. Furthermore, this is one of the four groups with many more missing cases in the validation samples than in the nonvalidation one (13.4% and 13.0% respectively). Corrections to AGI were made by 58 percent of the nonmissing applicants in this group, and the mean correction was to AGI $3,845 in one direction or the other. This was the largest amount for any group, resulting in a net underreporting of AGI of $1,230 (calculated among all non-missing applicants). On the other hand, this did not translate into very large discrepancies in SEI or award due to low income and assets.

Several possibilities exist. Since this group was more likely to over-claim than under-claim, the suspicion of some intentional misreporting and some failure to re-enter the system out of fear of getting caught must be entertained. If this is the case some other fields which are harder to validate, such as the absence of savings, could also be sources of error for this group. On the other hand, this could be a case of low income applicants trying to estimate their income carelessly and feeling intimidated by the validation requirement. The group should probably be selected for validation in the absence of remediation concerning estimation. Even if verification of tax returns

is required for all who estimate, at least part of this group should be validated with view toward possible errors in fields unrelated to tax returning or failure to report assets.

Group 3: dependents who estimated taxes, had SEI not over 400 and had savings. The presence of savings differentiates this group from the previous two. While the proportion of over-claimers among all applicants in this group is slightly higher, the proportion of misreporters among those not missing is lower than that of the previous group. (This is due to the absence of the extra 13 to 18 percent which failed to re-enter in the previous group when selected for validation.) Nevertheless, the percentage of over-claimers is inordinately high, reflecting the fact that the attempt to estimate income and taxes is, for the most part, likely to produce errors. The absence of additional missing cases when compared with the nonvalidation sample, and the fact that savings were reported suggests honest mistakes rather than fraud. The group should probably be selected for validation in the absence of remediation connected with the estimation of taxes.

Group 4: dependents who estimated taxes, had SEI over 400, taxes not over $2,000, no change in SEI prior to selection and were processed by May 31. This group has the highest proportion of misreporters (over-claimers and under-claimers combined) of any in the model. Fully 80 percent of non-missing applicants changed their SEI after selection, 56.4 percent of non-missing applicants doing so by 50 points or more. Among these applicants 71.4 percent corrected AGI and 63.7 percent, taxes paid. While this group misreports to a great degree, it is almost as likely to hurt itself as help itself by its misreporting. It even does some unsolicited correcting among nonvalidation applicants. This group should be selected for validation in the absence of other corrective action.

Group 5: dependents who estimated taxes, had SEI over 400, taxes not over $2,000, no change in SEI prior to selection and were processed after May 31. The total rate of misreporting for this group is slightly lower than for the previous one, it is still among the highest. There is also less unsolicited correcting among the nonvalidation applicants. It seems

less excusable to estimate taxes when one files after April 15 than when one files before, and it is with this sort of group that a change in procedures requiring verification with tax forms is in order. Failing some other sort of remediation, this group should be selected for validation.

Group 6: dependents who estimated taxes, had SEI over 400, taxes not over $2,000 and a change in SEI from a previous transaction. While still including more misreporters than the total average, this group is less error-prone than the two previous ones. Selection of this group is probably not indicated unless one were to select a large proportion of applicants for validation. More to the point is the fact that this group could be eliminated in most cases through corrective action. By the time an applicant receives an SER it is late enough in the year that he can verify his figures from his tax forms. The requirements that applicants do this and sign a statement to the effect prior to submitting corrections would completely eliminate applicants with this set of characteristics, and presumably reduce their misreporting.

Group 7: dependents who estimated taxes, had SEI from 401 to 1200, taxes over $2,000, both portions greater than 0 whose first choice is a university. This group has the largest percentage of over-claimers of all thirty-seven groups. It is a high income group to begin with (average AGI reported at selection is $25,536), and its members seem to be misreporting to their advantage. The average member of this group who re-enters the system gets $68 less after selection for validation than before. The average award is off by $99 (there are 9.6 percent under-claimers). In spite of this tendency to over-claim, this group also has the lowest proportion missing in both validation samples. AGI is corrected by 58 percent of its non-missing members and taxes paid by 59.1 percent. Validation seems to be in order for this group, unless an alternative means of handling the errors due to estimation of taxes is found.

Group 8: dependents who estimated taxes, had SEI from 401 to 1200, taxes over $2,000, both portions greater than 0, whose first choice institution is missing or not a university. This group has a lower rate

of over-claimers than the previous one, but still one of the highest rates of any group. Actually, a higher percentage understate AGI in this group (50%) than in any other group, but the amount of the discrepancies are not as large as in Groups 2 or 7. This group should also be validated.

Group 9: dependents who estimated taxes, had SEI over 1200, taxes over $2,000 and both portions greater than 0. This group had misreporting rates higher than the total sample, but lower than most applicants who estimated taxes. The fact that their award potential was low at selection limits the possible impact of validation. The group should not be validated, but corrective action with respect to estimation of taxes could be considered.

Group 10: dependents who estimated taxes, had SEI over 400, taxes over $2,000 and at least one portion 0 or blank. This is another group with a high rate of over-claimers. It has the highest net award difference from selection to criterion ($76 per applicant overpayment at selection). It was estimated in Chapter 4 that the difference between paying all recipients who fall in this group at selection and paying them after validation or other corrective action taken.

Group 11: dependents who did not estimate tax figures, filed through BEOG or PHEAA, did not own a house and chose a public institution or none at all. This group has very low rates of misreporting and a high rate of missing. The percentage missing is even higher among nonvalidation applicants, so validation could even be acting as an incentive to remain in the system. The group should not be validated nor is corrective action needed.

Group 12: dependents who did not estimate tax figures, filed through BEOG or PHEAA, did not own a house and their first choice school was private or proprietary. Similar to the previous group, but the percentage missing is lower. There is some evidence of validation keeping some applicants away, but the figure is small enough to be possibly due to sampling error. In any case applicants in this group should not be validated.

Group 13: dependents who did not estimate tax figures, filed through BEOG, house value was greater than 0 and chose a university or no institution. This group's misreporting pattern is close to average, and it should not be validated.

Group 14: dependent applicants who did not estimate taxes, filed through PHEAA, house value was greater than 0 and first choice was a university or no institution. This is another case of a group with high proportion of missing cases across both validation and nonvalidation samples. This is not surprising, since a large proportion of PHEAA applicants are interested in form of aid other than Basic Grants. Even though the rate of misreporting is higher than the previous three groups, this group does not merit validation.

Group 15: dependent applicants who did not estimate, filed through BEOG or PHEAA, house value was greater than 0 and listed a first choice institution other than a university. This is another group that resembles the total. Validation is not necessary.

Group 16: dependent applicants who did not estimate, filed through ACT or CSS, had SEI not over 200, total income not over $10,000 and NTI not over 50 percent of total income. While the rate of misreporting is below average for this group, there are a few cases misreporting in large scale, resulting in the largest net discrepancy in AGI between selection and criterion for groups other than the ones estimating tax figures. The percentage correcting AGI is not large (19.2%), but the average discrepancy for those correcting is $4,453! Selective validation or tighter edits, including verification of AGI, may well be in order.

Group 17: dependent applicants who did not estimate, filed through ACT or CSS, had SEI not over 200, total income not over $10,000 and NTI over 50 percent of total income. This group has a lower than average rate of misreporting, in spite of the fact that 26.4 percent of its applicants who re-enter the system underreported NTI. This field constitutes the bulk of the income for this group, the lowest in AGI. However, in spite of the fact that NTI is underreported by $473 on the average, this error seems not to affect SEI much because of low income and merits at most an edit, but not validation.

Group 18: dependents who did not estimate, filed through ACT or CSS, had SEI not over 200 and total income over $10,000. This group has a slightly below average rate of misreporting, but also a few cases who substantially underreport AGI. Selective validation of AGI or even an edit requesting verification might suffice here.

Group 19: dependents who did not estimate, filed through ACT or CSS, had SEI over 200 and either paid no taxes or left taxes blank. This group had a moderately high rate of misreporting, but it had more under-claimers than over-claimers. A good portion of this group's income is non-taxable and it had the second highest rate of NTI changes after selection. There is some evidence of applicants not re-entering the system as a result of selection. Again, the rate of misreporting is not among the highest, but some of the few who do misreport do so in a way that affects their SEI to a large degree. Among the 10.8 percent of the applicants who re-entered and corrected taxes the mean discrepancy was $1,333. An edit requesting verification of the tax field would be appropriate. Since this group receives on the average over $3,000 in Social Security benefits, these applicants may well find themselves in unusual circumstances meriting the suggestion that they seek help in filling out their form.

Group 20: dependents who did not estimate, filed through ACT or CSS had SEI from 201 to 600, paid taxes, but taxes were not over 5 percent of AGI. This group had a moderate rate of misreporting, close to the average. Furthermore, the impact of the misreporting that does exist in this group is rather low. No corrective action seems to be required for this group.

Group 21: dependents who did not estimate, filed through ACT or CSS, had SEI over 600, paid taxes, taxes were under 5 percent of AGI and assets were not over $30,000. This group had the largest proportion of under-claimers of any of the 37. Many of the under-claimers seem to have underreported taxes. The average applicant changing taxes paid after selection underreported by $1,107. This figure suggests the possibility of a digit being omitted. An edit pointing out the tax figure and a verification comment would be in order.

Group 22: dependents who did not estimate taxes, filed through ACT or CSS, had SEI over 600, paid taxes, had taxes not over 5 percent of AGI and assets over $30,000. Like the previous group, this group also has more under-claimers than over-claimers. The impact, however, is lower for this group. An edit or comment concerning taxes might be considered, but otherwise no remedial action is necessary.

Group 23: dependents who did not estimate taxes, filed through ACT or CSS, had SEI over 200 and taxes were 5 to 15 percent of AGI. This was the largest group, and its characteristics resemble those of the total sample. It is likely that further investigation, experimenting with other variables or other methods might discover further breakdowns of this group. To this end, the practice of random validation should be continued and the group given special attention in future attempts to refine the model. In the interim, no special remedial action seems indicated.

Group 24: dependents who did not estimate taxes, filed through ACT or CSS, had SEI over 200, taxes over 15 percent of AGI and AGI not over $25,000. This group had the greatest rate of over-claimers of any group which did not estimate taxes. Furthermore, the percentage of applicant who filed to re-enter the system after selection for validation was much greater than the corresponding percentage in these applicants to fail to re-enter the system over the percentage in the non-validation sample. By and large, overreporting of taxes was the most likely source of error, but underreporting AGI was also prevalent. Two possible conclusions can be formed. First, one may assume intent of fraud, and increased failure to re-enter out of fear of getting caught. A second possibility is that after validation many applicants in this group (whose mean selection SEI is over 1,000) become ineligible and do not bother to file the correction. In either case, members of this group should be validated.

Group 25: dependents who did not estimate taxes, filed through ACT or CSS, had SEI over 200, taxes over 15 percent of AGI and AGI over $25,000. The rate of over-claimers in this group was above average, but much lower than the previous one. Since awards were lower to begin with, the impact of misreporting was also lower. This group tended to increase the SEI even more after validation even though it was already high. In other words, over-claimers far outnumber under-claimers. The rates are not high enough to justify validation for all members of this group, but it would be well to keep an eye on results pertaining to it in the Task 2 analysis.

Group 26: independents who filed through BEOG or PHEAA with SEI = 0. This group has a very low rate of misreporting. Whatever errors are found in their applications, they are not ones that validation can identify. Perhaps quality control studies should check out for omissions on the part of these applicants, but for the time being no corrective action is needed.

Group 27: independents who filed through BEOG or PHEAA with SEI greater than 0. This group has a large proportion missing with a moderately high rate of misreporting among those re-entering the system. The net result of this misreporting is by a slight margin to the applicant's disadvantage. Validation is probably not justified for this group.

Group 28: independents who filed through ACT or CSS, had SEI = 0, exemptions not greater than one, were processed by April 30 and had AGI not over $2,000. This is another group with a low rate of misreporting. No corrective action is needed.

Group 29: independents who filed through ACT or CSS, had SEI = 0, exemptions not greater than one, were processed by April 30 and had AGI over $2,000. This is a group very likely to have been affected by the SEI formula change that went into effect May 1. Their SEI under the new formula is zero, but their old SEI probably was not. Fully 75.5 percent

of them are missing in the nonvalidation sample, but validation seems to have acted as an incentive for them to remain in the system. No corrective action is needed, but the group should be observed in a year without an SEI formula change.

Group 30: independents who filed through ACT or CSS, SEI = 0, exemptions not greater than one, processed May 1 to July 31. This is another low misreporting group for which no action is needed.

Group 31: independents who filed through ACT or CSS, SEI = 0, exemptions not greater than one, processed after July 31. This group has low misreporting rates, but many more missing in the validation samples than in the nonvalidation sample. It is a good indicator that validation can delay re-entry. No corrective action is necessary.

Group 32: independents who filed through ACT or CSS, had SEI = 0, exemptions greater than 1 and year in school 1 or blank. This is another low error group. No corrective action is needed.

Group 33: independents who filed through ACT or CSS, had SEI = 0, exemptions greater than 1 and year in school greater than 1. This group differs from the previous only in having fewer missing cases. No corrective action is needed.

Group 34: independents who filed through ACT or CSS, had SEI greater than 0, were processed by April 30, paid taxes not over $500, and signed their application by February 28 or left the date blank. This group combines a high missing rate with high rates of misreporting. Since the group could have been partially affected by the SEI formula change, it is hard to judge the results. One hesitates to recommend validation of this or the next two groups based on observations from a year with an SEI formula change and a confusing systems generated transaction. Yet at the very least further study should be done.

Group 35: independents who filed through ACT or CSS, had SEI greater than 0, were processed by April 30, had taxes not over $500 and signed the application after February 28. This group had fewer misreporters than the previous one, but the rate is still higher than the average for independents.

Group 36: independents who filed through ACT, had SEI greater than 0, were processed by April 30, and had taxes over $500. Over two thirds of these applicants failed to re-enter the system, but most of those who did were misreporters. Validation of just this group among independents might serve to obtain a better picture. However, results from this study could be confounded by the SEI formula change, a fact that makes suspicious any results separating applicants processed by April 30 (THAID was provided cut-off dates for the end of each month) from all others. On the other hand, the results might have had more to do with the income tax filing date than with the SEI formula change.

Group 37: independents who filed through ACT or CSS, had SEI greater than 0 and were processed after April 30. This group had fewer missing cases than the preceding ones, but still a rate of misreporting higher than most independents. The rate is not high enough to warrant validation.

Chapter 8 will coordinate the various comments into recommendations for use of the findings and will compare them with the PEC.

8

RECOMMENDATIONS AND EFFECTIVENESS OF THE MODEL

This chapter presents various recommendations for the use of the error-prone model, and compares its effectiveness with the PEC. The first section presents a strategy for using the model to select applicants for validation. An alternative approach related to estimation of taxes is discussed in the second section, followed by a section on validation of independents. A comparison of the relative effectiveness of the EPM and the PEC follows. The chapter concludes with recommendations for further research and for program improvement.

8.1 A Validation Selection Strategy in the Absence of Other Changes

This section presents a strategy for selecting applicants for validation based on the model, assuming the model is the only information available and validation the only possible action to be taken. If OSFA makes dramatic changes to the delivery system, the model would not be valid. Similarly, if validation procedures are changed, the model likewise would have to be adjusted. The strategy simply consists of selecting all applicants from groups 2, 3, 4, 5, 7, 8, 10, and 24. These groups include 11 percent of all applicants in the random validation file. This combined group includes 23.3 percent over-claimers, ten percent under-claimers, and 26.2 percent missing. Of applicants who re-enter the system, 45.5 percent of members of these groups are

misreporters. The group would include 28.7 percent of all over-claimers (or at least of those who filed by August 31). An additional one percent of applicants could be selected on a random basis to ensure continued investigation, for a total of 12 percent selected, accounting for 29.5 percent of all over-claimers.

As will be noted, no independents are included among the eight groups recommended for selection. This is due to the much lower error rates for independents found in this study, which is probably somewhat due to the problem presented by the change in the SEI formula.

This validation procedure would be more effective than the PEC taken collectively. The PEC validation file has only 13.6 percent over-claimers and 29.2 percent of those who re-enter the system misreporting. Some individual PEC are more effective, and the possibility of combining the two will be explored in a later section.

8.2 Estimation of Taxes, an Alternative Remedy

Seven out of the eight groups recommended for validation in the previous section had estimated tax figures. Estimated tax figures by dependents accounts for over 30 percent of all misreporters. A good proportion of these errors could be eliminated through various policy changes. Among these are:

- Instruct applicants to obtain a 1040/1040A form and complete it, obtaining help from a tax service if necessary, if they have not filed a tax return by the time they complete their applications, and to sign a statement to that effect.
- Require that a copy of 1040/1040A returns be enclosed with application (possibly schedules filed with the 1040).
- Require that financial aid administrators validate tax return only, for applicants who estimated taxes.
- Conduct verification of income and taxes by mail for applicants who estimated (this allows early filing, but requires corrections if error).
- Provide a 1040 form and directions along with applications and instruct that it must be completed before the application is completed.

Implementation of one or more of these recommendations would probably reduce error considerably. The change in policy, however, would require a revision of the model since estimation of taxes would no longer provide the information it implied in 1979-1980. Many misreporters who may have used estimation as an excuse for over-claiming, might fall into other groups, but all in all, error would almost certainly be reduced.

8.3 Validation of Independents

Two issues related to misreporting patterns of independents need to be discussed. First is the problem of the SEI formula change. This study must remain inconclusive as to the behavior of independents whose applications were first processed by May 8, 1979, due to circumstances which are not likely to recoccur in subsequent years. Thus, groups such as 34, 35, and 36 may or may not merit validation. The Task 2 analysis may be able to answer some of these questions, since more complete data will be available for that analysis.

The second issue is that the SEI is the best predictor of misreporting (with MDE source predicting re-entry). Applicants with SEI=0 are not found to be misreporting frequently through current validation procedures. This may be because the less the applicant reports, the fewer things the FAA is able to validate. In addition, it is very difficult to validate dependency status, and the most effective way for an applicant to misreport is to improperly claim independent status. Thus, if independent applicants were to be selected for validation under present validation procedures, groups 27 and 34 to 37 would be selected. It would be less effective, however, to select these groups than to concentrate on the ten dependent groups mentioned in section 6.1. On the other hand, if a more effective way of validating low income applicants can be devised, then the present model should not be used to select applicants for that kind of validation.

8.4 The PEC Versus the EPM

Various approaches were used to compare the effectiveness of the PEC with that of the model. One approach was to code members of the random validation group who would have met PEC by which PEC they met. A multiple linear regression with absolute award difference as the dependent variable and the EPM and PEC as predictors was conducted for applicants who re-entered the system. The EPM accounted for 4.3 percent of the variance of award difference. The PEC accounted for 1.5 percent of the variance. Combining the two accounted for 5.5 percent of the variance of the dependent variable. The net result is that the EPM improves the PEC much more than the PEC improves the EPM.

Looking at the PEC validation file, however, it became apparent that some PEC were working well and others not at all. PEC that worked well were A6, A10, A11, A14, and A15. Each of these yielded over 20 percent over-claimers, with A6 working better than any group in the EPM (37% over-claimers). This, however, includes an over-representation of applicants which met other PEC. Among applicants in the PEC file which met A6, 25 percent belong to one of the eight groups suggested for selection in section 6.1. PEC A6 involved a complex formula which was not used as a variable, so it does not appear directly in the model. The other PEC involved small percentages of the population and would not have appeared had they been included.

It would probably be desirable to incorporate these five PEC into the selection criterion. They should be seen as an additional check along with the error-prone model. On the other hand, the PEC other than these five are probably unnecessary if the EPM is being used.

8.5 Further Research Needed

The present model is a first step which needs to be elaborated and refined. Its validity in subsequent years should be tested, as should its validity for applicants selected after August 31. The model is only appropriate for current validation procedures, and should be refined as new methods of quality assurance are developed.

Quality control studies should investigate why applicants make mistakes. Reasons should then be broken down by EPM group and appropriate edits generated or validation indicated when a pattern seems clear.

Subsequent investigations may subdivide some groups and merge others. Policy changes such as requiring tax forms or inclusion of tighter edits may make the model obsolete. Matching the model with data obtained from the IRS analysis will be useful in obtaining "true" (vs. FAA-verified), income, as well as obtaining income for applicants classified as missing in this study.

New variables presently available from the 1979-1980 data files should be tried, particularly with a view to breaking up the larger groups. Variables not presently available, but which could be acquired in a QC study should be used in developing new models, so that those which become successful predictors can be incorporated into future application forms. A top priority should be identifying those cases of serious misreporting that are scattered among groups with low error rates.

8.6 General Recommendations and Alternatives Related to the EPM

Several suggested improvements for the use of the results presented in Part II for quality assurance follow:

- Select groups 2, 3, 4, 5, 7, 8, 10, and 24 for validation
- Supplement EPM with PEC A6, A10, A11, A14, and A15.
- Attack the error rate due to estimation. Several suggestions have been made in section 8.2. Consideration should be given to wider use of form 1040 for quality assurance.
- If independents are to be selected under present validation practices, select only those with SEI greater than 0.
- List an applicant's group number in the SER. Provide school counselors and Financial Aid Administrators with profiles of each group (but not with validation criteria) as an aid to identifying possible sources of error and to helping under-claimers receive the amount they are entitled to.

APPENDIX A

COMPUTER COMMENTS

APPENDIX A

CRITICAL FIELD COMMENTS FOR NONSUPPLEMENTAL APPLICANTS

Rejection			Assumption	Other	
10	131	300	4	9	246
12	182	301	5	29	253
13	202	302	6	50	254
18	212	303	7	51	255
23	234	306	11	52	257
24	236	307	14	66	258
27	238	308	15	68	260
28	247	309	17	69	295
30	251	310	19	71	296
33	256	311	20	72	
35	259	312	21	73	
41	268	313	22	75	
42	269	314	25	76	
43	270	315	26	77	
44	271	316	31	78	
47	273		36	79	
48	275		37	80	
58	277		38	81	
70	279		40	82	
98	280		107	84	
99	281		242	85	
106	282			97	
128	283			108	
130	284			109	
135	287			116	
138	288			117	
139	289			119	
140	290			120	
148	297			121	
179	298			122	
180	299			123	
				124	
				125	
				132	
				134	
				200	
				204	
				213	
				240	
				244	

APPENDIX B
ASSUMPTIONS FOR COMPUTING ELIGIBILITY INDEX WHEN APPLICANT IS REJECTED

ASSUMPTIONS FOR COMPUTING ELIGIBILITY INDEX WHEN APPLICANT IS REJECTED

Dependency Status

If model is missing:

if side d-or-e is = "D" then applicant is dependent

if side d-or-e is = "E" then applicant is independent

If side d-or-e is missing:

if model-ques is all "no" then applicant is independent; otherwise applicant is dependent.

Agi and Portions

(a) if agi is blank and 2 portions present, compute agi by adding the 2 portions;

(b) if agi is blank and 1 portion present, compute agi by using the 1 portion;

(c) if agi is blank and no portions present, assume agi and both portions to be zero;

(d) if agi is present and both portions blank, assume 1 portion.

Non-Taxable Income (nt-total)

(a) if nt-total is blank and nt (social security benefits) is blank and nt (other) is blank, assume nt-total to be zero.

(b) if nt-total is blank and one nt portion is present, assume nt-total to be equal to the present nt portion.

(c) if nt total is blank and both nt portions present, assume nt total to be sum of two portions.

Unusual Expenses (ue)

(a) if medical and dental expenses is present and casualty theft loss is present add both fields to give ue.

(b) if either field is missing, use the field present as ue.

(c) if both fields are missing, assume ue to be zero.

Unreimbursed Tuition

If missing, assume zero.

Assets/Cash and Savings

If any field is missing, assume zero for that field.

Taxes Paid

If taxes paid is missing, use compute formula with assumed values for related fields (agi, hh-size, etc.) if necessary.

Marital Status/Household Size

(a) Marital Status and Household Size Missing:

(a-1) Dependents: marital status - married
household size - 3

(a-2) Independents: marital status - single
household size - 1

(b) Marital Status Missing/Household size present:

(b-1) Dependent, hh-size greater than 2:
assume married

(b-2) Dependent, all other hh-size:
assume divorced

(b-3) Independent, hh-size =1:
assume single

(b-4) Independent, all other hh-size:
assume married.

Household Size (hh-size) (Only if hh-size has not been assumed along with marital status).

(a) If hh-size is missing and exemptions is present, use exemptions.

(b) If exemptions is missing:

If applicant is dependent and parent's' marital stattus is married, assume 3.

If applicant is dependent and parent's' marital status is divorced, single, widowed or separated, assume 2.

If applicant is dependent and parents are deceased, assume 1.

If independent and married, assume 2.

If independent and single/widowed/divorced/separated, assume 1.

Veteran's Benefits

(a) If amount and months are missing, assume zero.

(b) If amount is present but months are missing, assume months to be 9.

(c) If amount is missing and months is present:

if hh-size is zero, assume 249
if hh-size is one, assume 296
if hh-size is two, assume 338
if hh-size is 3 or more, vet-ben = 338+
(26 times number in hh greater than 2).

APPENDIX C

PRE-ESTABLISHED CRITERIA

Pre-established Criteria

PEC A: Applicants who establish eligibility after rejection of a transaction through correction where the rejection reason meets one or more of the following categories:

- A-1: Missing information for household size and marital status
- A-2: Portions and AGI inconsistent
- A-3: Unusual expenses inconsistent with income or greater than $5,000
- A-4: Income negative, zero or less than $51
- A-5: Unreimbursed tuition greater than 35% of income
- A-6: Income and taxes paid inconsistent
- A-7: Financial data missing
- A-8: Financial data missing
- A-9: Applicant's resources reduced
- A-10: No S.S. match and applicant verifies
- A-11: No S.S. match and applicant changes
- A-12: V.A. blank or zero and applicant verifies
- A-13: V.A. blank or zero and applicant changes
- A-14: V.A. greater than zero and less than $131 and applicant verifies
- A-15: V.A. greater than zero and less than $131 and applicant changes
- A-16: Combination of an A criterion with B1
- A-17: Combination of an A criterion with B2
- A-18: Combination of an A criterion with C1
- A-19: Combination of an A criterion with C2
- A-20: Combination of an A criterion with D1
- A-21: Combination of an A criterion with D2

PEC B: Applicants who establish eligibility with one correction resulting in a large eligibility index change.

- B-1: Large EI change
- B-2: Larger EI change
- B-3: Combination of a B criterion with C1
- B-4: Combination of a B criterion with C2
- B-5: Combination of a B criterion with D1
- B-6: Combination of a B criterion with D2

PEC C: Applicants who reduce their highest EI in one continuous official string of transactions.

- C-1: Large EI change
- C-2: Larger EI change
- C-3: Combination of a C criterion with D1
- C-4: Combination of a C criterion with D2

PEC D: Number in postsecondary education (PHE) and tax filing status.

- D-1: PHE greater than 3
- D-2: Tax return figures are estimated
- D-3: Combination of D1 and D2

APPENDIX D

GENERAL METHODOLOGY FOR ERROR-PRONE MODEL

Error-Prone Modeling Techniques

The development of adequate error-prone models has become an important concern for various agencies. The need to determine which cases are most likely to be misreporting so that correction action-procedures can be instituted is likely to be present for any program engaged in the disbursement of public funds on the basis of stated need and qualifying conditions.

Three major approaches to error-prone modeling have been used by various state and federal agencies. Each has its advantages and disadvantages and each is best suited to different kinds of situations.

The first method, used by the Welfare Departments of South Carolina and the District of Columbia, uses discriminant analysis to obtain a formula which assigns a score to each case. The higher the score, the more likely it would be that the applicant is misreporting. Thus if the agency wanted to select applicants most likely to misreport, it would simply select those applicants to whom the formula assigned the highest scores.

The major drawback of this method is that it ordinarily assumes that a variable will affect all applicants in the same way. If, for example, it turns out that estimating taxes is significant for dependent but not independent applicants, discriminant analysis will fail to take this into account. Thus it could easily fail to detect some important combintions of variables which could predict error-proneness. In addition, discriminant analysis would not point to specific areas where an applicant is likely to be misreporting. Since each applicant receives a single score one cannot distinguish those who misreported AGI from those misreporting rates paid. Of course, separate analyses could be conducted to predict misreporting for each specific field, but this method would lack parsimony and would be difficult to interpret.

The second approach, used by the State of New Hampshire to identify error-prone cases in its MEDICAID program, attempts to define a single group most likely to exhibit a high degree of misreporting. Depending on the size of the group, every member or a certain percentage of this group

(and this group only) would be validated. Where only a small proportion of all the cases can be validated, and the principal objective is to maximize the savings in actual disbursements from the cases actually validated, this method can be very effective. On the other hand, this approach is likely to overlook groups whose error rate might approach that of the selected group, and which might require a different type of corrective measure. The BEOG program, with the use of edits in addition to validation, a high drop-out rate among applicants selected for validation and the dual concern for deterrence as well as savings in disbursements to validated applicants, requires a different approach.

The third approach has been used by the State of West Virginia Aid to Families with Dependent Children program and by the Supplemental Security Income program of the Social Security Administration. It essentially divides the applicant population into mutually exclusive groups which differ as much as possible from each other in either the mean dicrepancy in expected disbursement or in the rate or type of misreporting. This method has the potential to describe each group separately in terms of type of error, and thus to prescribe different types of corrective action for each. It has the further advantages of taking into account effects which apply to only part of the population, and of producing results which can be expressed in simple terms. This method, sometimes called classification analysis, sequential structure search or automatic interaction detection (AID) is the one which will be used in this study.

Overview of the Sequential Search Approach

The term sequential structure search is the more generic term for a conceptual method of exploratory analysis designed to discover nonlinear combinations among many variables which best predict a single dependent variable. The term Automatic Interaction Detector (AID) is often used synonymously with sequential structure search, but is often used more precisely to describe the implementation of this concept by the Institute of Social Research (ISR) at the University of Michigan. ISR has developed two programs which will conduct this type of analysis: AID3 which accepts continuous dependent variables, and THAID which accepts categorical dependent variables. This software was used by the

Supplemental Security Income program in their development of error-prone profiles. In addition, the West Virginia AFDC program used a software package closely related to THAID. In the subsequent discussion, the term AID will refer to the general technique rather than to a specific program or package.

Techniques such as discriminant analysis or multiple linear regression make the assumption that a given effect will apply equally to all members of the population. The result of either of these techniques is an equation which is meant to predict the dependent variable for every member of the population. AID, on the other hand, does not make the assumption that a predictor variable will affect the dependent variable in the same way for all cases. Instead, AID starts by breaking up the sample into two subgroups selecting that split which produces groups that differ from each other on the dependent variable as much as possible. Each subgroup is then split separately, allowing for different predictor variables to split different previously-formed subgroups.

Interaction effects occur when a variable predicts differently for one group than for another. Ordinarily, discriminant analysis and multiple linear regression do not take interaction effects into account. AID specifically identifies groups (using various combinations of variables) which will differ as much as possible on some criterion variable. Thus AID will be able to identify error-prone cases in instances where, for example, low taxes are an indication of error-proneness among high income applicants, but not for low income applicants. Linear models are oblivious to such relationships.

AID accepts one dependent variable, which may be categorical (such as type of applicant) or quantitative (such as discrepancy in expected disbursement, expressed in dollars). Predictor variables can be monotonic (where the sample or any subgroup can only be subdivided into high and low groups based on some cut-off point) or free (used for categorical, as opposed to quantitative, predictors where any combination of values can be used to split the groups). In either case predictor variables must be coded in terms of a small number of possible values (no greater than 10 for THAID).

Out of the many possible splits defined by predictors, AID selects the one which will divide the sample into two groups as different from each other as possible (the operational definition of "as different from each other as possible" is given in Chapter 5). The process then continues for each of the two groups into which the original sample was split. When a group becomes very homogeneous, cannot be further split using the predictor variables available, or would yield subgroups under a certain size if it were split, then the process is complete and it becomes one of the groups defined by the model. If an applicant group is both large and heterogeneous, it would be an indication that some additional predictors should be sought and included in the analysis.

Because AID investigates many possible combinations of variables, at times it produces results which are specific to a given sample. Two questions may be asked pertaining to the groups which emerge from an AID analysis: (a) Do the groups have the same characteristics in the population as they appear to have in the sample? and (b) do the groups constitute an optimal classification of the population if one is trying to predict the dependent variable? In order to answer either question one needs to use a second sample randomly drawn from the same population. The first question can be answered by checking whether the subgroups produced by AID from the first sample have similar characteristics in the second sample. The second question requires that a separate analysis be conducted for the second sample. It is quite possible that one would obtain a different solution if variables which are highly interrelated are used (this is similar to the problem of multicollinearity in multiple linear regression). The question of whether a given solution is the best possible, however, is of secondary importance to whether the classification which emerges is effective in predicting error-proneness.

The THAID Program

While AID 3 is the most popular sequential search program, it does not handle nominal scale dependent variables which achieve more than two different values. THAID, on the other hand, is precisely designed to

handle nominal scale dependent variables which achieve up to ten different values. THAID accepts up to 40 predictor variables, each of which can also achieve up to ten values.

THAID, like AID3 splits each sample or subsample into two subgroups in such a way that the two subgroups are as different as possible in terms of the dependent variable. Two different statistics--theta and delta--are available as operational definitions of "as different as possible." The theta statistic (from which the program gets its name) allows only splits where the two subgroups created by the split have different modes with respect to the dependent variable. In other words, the value of the dependent variable appearing most frequently in the first subgroup must be different from the value appearing most frequently in the second subgroup. THAID selects at each step that split which maximizes the value of theta, where

$$\text{Theta} = \frac{M_1 + M_2}{N_1 + N_2}$$

M_1 and M_2 represent the frequency of the mode in each subgroup (i.e., the number of cases attaining the value most frequently attained by the dependent variable in each subgroup) and N_1 and N_2 represent the total number of cases in each subgroup.

Theta has the advantage of producing subgroups with different modes if the data allow it, an ideal situation when the purpose is selecting cases which achieve a given value on the dependent variable (in our case, applicants who are misreporting to their advantage). The disadvantage is that the data seldom allow it when one has a lopsided distribution, and in this study over-claimers and under-claimers combined do not reach 15 percent of the sample.

The alternative is the delta statistic, which selects splits which are as different from each other as possible, taking into account every value of the dependent variable, not just the mode. When THAID splits the sample or a subsample into two subgroups it calculates the delta statistic for each split.

$$\text{Delta} = \frac{N_1 \sum_{j=1}^{G} |P_j - p_{1j}| + N_2 \sum_{j=1}^{G} |P_j - p_{2j}|}{2\left(N - \frac{\sum N_j^2}{N}\right)}$$

Where

N_1 = frequency in the first subgroup

N_2 = frequency in the second subgroup

P_j = proportion of cases in unsplit group falling on the j^{th} code of the dependent variable

p_{ij} = proportion of cases in i^{it} split subgroup falling on the j^{th} code of the dependent variable

N = frequency in unsplit group

N_j = frequency in unsplit group in j^{th} code

G = number of possible values in the dependent variable.

THAID choses that split which can be defined by a predictor variable which maximizes delta.

THAID produces four iterations, and therefore up to 16 groups can result from one run. However, one can also restrict the cases which will be included in a given run, so that a group which was defined by one run on the fourth iteration can become the entire sample for a subsequent run. Thus, in practice any number of iterations can be produced.

The option to restrict the sample can also be used to force a first split or subsequent splits on theoretical grounds. This option was used to start off by splitting dependents from independents. Other forced splits were also tried, but not used in the final model.

The output provided by THAID can include upon request information on the best split that could have been produced by each of the predictors at each step, and the value of theta or delta for each predictor in each instance. this information can be used to prepare alternate runs with forced splits or to decide which predictors to delete in subsequent runs.

Other options available in the THAID program include:

- a weight function which can be used with a stratified sample or if one wishes a value of the dependent variable to appear to be the mode if it exceeds a certain percentage of the cases.
- an option to limit the minimum size a group can achieve.
- an option to set a minimum value for theta or delta before a split can be produced.

Limitations of the Approach

One of the major limitations of this approach to error-prone modeling is that one can never be certain of having produced the best possible solution. A different investigator, using different samples, different variables, or forcing a different first split could well come up with a totally different model. What could have been an effective split at the second iteration may never appear in light of a different split at the first.

THAID has a tendency to select splits that come close to dividing a sample or subsample in an uneven division. For this reason criteria which define only a very small percentage of the population would be unlikely to appear in the model. The one tenth of one percent of the applicants defined by a given PEC could be expected to be scattered among the thirty-seven groups resulting from the study. For this reason the model should not be used to the exclusion of other criteria that have also proven effective.

Limitations of the study (as opposed to the THAID method in general) will be discussed in Chapter 6. Chapter 3 will discuss the specific way in which THAID was applied in the present investigation.

APPENDIX E

STEP-BY-STEP DEVELOPMENT OF THE ERROR-PRONE MODEL

This appendix describes the sequential search procedure used in developing the model. Each derived model and subsequent refinement is briefly explained. The very first THAID runs were conducted on the working sample using only variables which could be obtained directly from the SER, plus the variable "number of prior transactions". While these runs were performed on dependents and independents separately, the independent runs produced erroneous results since the systems generated transaction and the change in EI formula had not been accounted for.

These runs used the delta statistic placed at 100 the minimum membership a group was allowed to have. After the first run produced sixteen groups (for dependents only) each group was in turn submitted to further iterations. The cross-validation attempt using the replication sample revealed some of the splits to be idiosyncratic (i.e., specific to the sample), and in the end twenty groups were retained.

A split was considered to be replicated if a similar division was obtained in the replication sample; this was determined by comparing the resulting four subgroups--two in the working sample with two in the replication sample. In other words, if W1 and W2 are the two subgroups resulting from a split in the working sample and R1 and R2 the corresponding groups in the replication sample, W1 must be more similar to R1 than to W2 or R2, and so on for the four groups. The decision on similarity, however, was made on an intuitive basis, and some legitimate, but trivial, splits were at times discarded if the resulting groups were small.

The very first model, derived exclusively for dependent applicants, was a reasonably adequate model except for a few conceptually inappropriate characteristics. One of the defining variables was non-taxable income other than Social Security. The program separated those who left this field blank from those who did not, but in doing so it was merely selecting PHEAA applicants (who did not have a place to include this field in their applications) and BEOG applicants who left the field blank, from the rest of the BEOG applicants.

A second inadequacy of the model was that taxes paid as a proportion of AGI was coded as a free variable, a THAID option which allows splits other than "high" vs. "low." This was very effective since the group with the middle range is less likely to misreport than those at the two extremes. The problem is that some subgroup consisted primarily of members of one extreme (low taxes with respect to AGI, for example) with a few stragglers from the other extreme. This made the definition conceptually awkward and reduced the homogeneity of the groups.

Variables which were not effective at any step of the process or which if effective failed to replicate were removed. The coding for some variables was changed. NTI replaced Social Security Benefits and Non-taxable-other. Taxes as a proportion of AGI was refined and changed so as to allow only "high" vs. "low" splits. New ratios were created in view of the success of "taxes as a proportion of AGI". Corrections related variables were also created.

At this point a number of experimental runs were conducted using only data from the dependent applicants. Variables were added, and in some cases subsequently eliminated. Splits based on theoretical expectation were forced on the data. The theta statistic was used. Cases were weighred based on their score on the dependent variable (TYPE) so as to increase the probability of finding small groups with large percentages of over-claimers and under-claimers, and analyzed using both the theta and the delta statistic. The roles of the working and replication samples were switched.

It soon became apparent that theta as a splitting criterion was not appropriate regardless of how one weighted the data. It also became apparent that three variables were competing for a role at the very first split of the dependent subsamples. Very slight changes in the data or the procedures allowed one or another of these variables to create the first split, and depending on which variable was chosen by the program, the subsequent splits produced very different results. These three variables were tax filing status (separating those who estimated from all others, MDE source (separating BEOG and PHEAA applicants from ACT and

CSS) and SEI (with various cutoff points at the lower end of the spectrum). When no weights were used, the first two of these variables produced very similar deltas.

When tax filing status was used as the first splitting variable, MDE source proved effective for those who did not estimate taxes (used tax form, did not have to file or did not answer the question). When MDE source was selected to create the first split, tax filing status was effective for ACT and CSS applicants, but not for BEOG and PHEAA. Eligibility index appeared frequently, but not first, in the unweighted runs, while often creating the first split with certain combinations of weights. Weighted runs often discovered small groups with high percentages of misreporters, but were also more likely to yield results which did not replicate.

Corrections variables in general did not predict well. This was thought to be because it is difficult to define corrections variables which apply to a large proportion of applicants and are at the same time effective. The desirability of isolating cases where corrections history might make a difference became apparent.

Forty variables were identified as offering the best possible chances for success. Three differfent procedures were implemented simultaneously, using the working sample. The first procedure was a simple, unweighted THAID analysis, forcing the first split on dependency status, using the delta statistic and setting at 50 the minimum number of cases in any group resulting from a split. The second was identical, except that a second forced split, separating first transactions from those with prior transactions was implemented. The third procedure used a weighting scheme designed to produce equal weighted frequencies for all four applicants types for each dependency status. Since group size had to be determined using weighted frequencies, they were set so group sizes would average 100, but could be smaller for groups with large proportions of misreporters. The delta statistic was also used here.

The simple unweighted model proved to be the most effective, after incorporation of a few additional divisions suggested by the early model. The forced split on prior transactions model yielded more groups,

and did use some of the corrections variables. It also left some large groups unsplit to a greater extent than the simple model, did not identify groups with results as extreme as those of the simple model, and was less parsimonious. The weignted model identified a few more extreme groups, but left many large groups unsplit.

The simple model was selected and further investigated. Proportions of misreporters for nonvalidation applicants were calculated for each group. Mean values of the critical fields, SEI and scheduled award on the criterion and selection transactions were calculated for each group using the working sample (computed with and without absolute values). Percentage of non-missing applicants correcting each critical field in each direction were also calculated for each group. Thus a profile of each of the groups defined by the model was obtained.

APPENDIX F

GLOSSARY OF TERMS

GLOSSARY OF TERMS

ACT:	American College Testing Program. ACT's Family Financial Statement (FFS) is one of four applications with which a student may apply for a Basic Grant. (See MDE for a list of other sources.)
AGI:	Adjusted gross income. AGI is an application item which refers to wages, salaries, tips, farm income, dividends, interest, and business income, as reported to the Internal Revenue Service.
AR:	Applicant's resources. AR is an application item which refers to the total of the dependent applicant's (and spouse's) cash; savings and checking accounts, and other assets.
Assumption Edit:	One type of the processing system's series of computerized edits designed to check for missing information, as well as the logic and consistency of all data provided. With this edit, the processing system assumes a value for a missing or apparently incorrect application item based on other provided data.
BEOG:	Basic Educational Opportunity Grant. BEOG's application form is one of four forms with which a student may file for a Basic Grant. (See MDE for a list of the other sources.)
Comment:	An instructional message on the SER. Many comments advise the student to review the application and take further action, if necessary. The computerized comments are triggered by the processing system edits.
Composite Index:	An indicator used in this study to measure applicant corrections behavior. The composite index is calculated by multiplying the percentage of applicants making corrections by the effective SEI change.
Correction:	Following the initial application, a student may change any or all of his or her application information. A correction may be in response to an edit or the validation process, or it may be done at the student's initiative. For analytical purposes, corrections are divided into ones occuring pre- and post-selection/eligibility.
Critical fields:	The 14 application items used in this study. These items are: AGI, AR, HS, Model, MS, NA, NTI, PHE, Portions, Tax Filing Status, TP, UE, UT, and VEB.

CSS:	College Scholarship Service. CSS's Financial Aid Form (FAF) is one of four applications with which a student may apply for a Basic Grant. (See MDE for a list of other sources.)
Edit:	One feature of the application processing system designed to minimize the number of student eligibility determinations made on the basis of invalid, inaccurate or incomplete data. The edits screen the applications by checking for missing information and the logic and consistency of all data provided.
Effective SEI Change:	The actual point change resulting from a correction if SEIs above 1600 are set at 1600. For example, the absolute change for an applicant whose SEI goes from 500 to 2200 is 1700; however, the effective change is 1100, or 1600 minus 500.
Eligible:	The status of an applicant who is qualified to receive a Basic Grant. Eligible applicants have an SEI from 0 to 1600.
EPM:	Error-Prone Model. The result of a procedure to identify applicants who are likely to misreport on their Basic Grant applications.
Expected Disbursement:	The amount of award an applicant is due to receive, taking into account enrollment status, cost of education at the institution indicated on the application, and SEI. The applicant's award is referred to as "expected" because the data on the recipient file is not validated until after the end of the academic year. However, it is expected that expected disbursements are reasonable predictors for actual disbursements.
Expected To Be Paid:	Refers to applicants appearing of the recipient file.
FAA:	Financial Aid Administrator.
HS:	Household size. HS is an application item which refers to the number in the household for whom the parents (dependents) or the applicant (independents) planned to provide more than half the support during the 1979-80 academic year.
Ineligible:	The status of an applicant with an SEI greater than 1600 who is not qualified to receive a Basic Grant.

Model: The dependency status--independent or dependent--of the applicant. An applicant is a dependent if he or she answered "yes" to one or more of the following application questions:

> Did or will the student live with the parents for six weeks or more in 1978, 1979, or 1980?
>
> Did or will the parents claim the student as a tax exemption in 1978, 1979, or 1980?
>
> Did or will the student receive $750 or more worth of assistance from the parents in 1978,-1979 or 1980?

MS: Marital status. MS is an application item which refers to whether the applicant (independents) or the parents (dependents) are single, married, divorced, separated or widowed.

MDE: Multiple Data Entry. Process by which an individual can apply for a Basic Grant using any one of the following four application forms:

- CSS's Financial Aid Form (FAF)
- ACT's Family Financial Statement (FFS)
- PHEAA's application
- BEOG application

NA: Net assets. NA is calculated from several application items. It equals the sum of the estimated market value of the applicant's (independent) or parents' (dependent) home, real estate, investments, business, farm, and checking and savings accounts minus the liabilities on these items.

Nonvalidation: Applicants who were not selected by the processing system to present certain documents to their financial aid administrators which confirm the accuracy of the information on their application form.

Not Expected To Be Paid: Refers to applicants not on the recipient file.

NTI: Nontaxable income. NTI is the sum of two application items which refer to the amount of social security benefits, child support, welfare, unemployment compensation, veteran's benefits (excluding veteran's educational benefits) and other income not gained through employment and not subject to U.S. income tax.

PEC: Pre-established criteria. The PEC are standards used to identify for validation those students most likely to misreport information on their application. The criteria, based on prior analyses of misreporters, set up categories of applicants with questionable information on their application or suspicious corrections behavior, and target a certain number of applicants in each category for validation. (See Appendix D)

PHE: . An application item which refers to the number in the applicant's (independent) or parents' (dependent) household which planned to be enrolled at least half-time in a postsecondary educational institution during the 1979-1980 academic year.

PHEAA: Pennsylvania Higher Education Assistance Agency. PHEAA's application form is one of four forms with which a student may file for a Basic Grant. (See MDE for a list of the other sources.)

Portions: The sum of two application items which refer to the income earned through employment (wages, salaries and tips) of both the applicant and spouse (independent) or the mother and father (dependent).

Pre-/Post-Eligibility: Refers to specific transactions of nonvalidation applicants. Post-eligibility transactions are subsequent to the transaction when the applicant achieved eligible status. Pre-eligibility transactions include both the transactions prior to eligibility and the transaction when the applicant achieved eligible status.

Pre-/Post-Selection: Refers to specific transactions of validation applicants. Post-selection transactions are subsequent to the transaction when the applicant is selected for validation. Pre-selection transactions includes the selection transaction and all previous transactions.

Random: Refers to the small portion of validation applicants that were not selected for validation as a result of pre-established criteria.

Rejected: The status of an applicant for whom an eligibility determination cannot be made because of missing information or suspiciously inconsistent application items.

Rejection Edit: The most restrictive of the processing system's computerized edits. Following the receipt of a rejection edit, the applicant must provide missing information or verify or correct existing data before the SEI can be generated.

Reject Reason Code:	Each of the 43 reject edits is identified with a code from alpha A to Z to numeric 1 to 17.
SEI:	Student Eligibility Index. A numeric indicator of the applicant's financial strength based on information provided on the application. When the SEI is in the eligible range--from 0 to 1600--it is combined with the applicant's educational cost to determine the size of the applicant's grant.
SER:	Student Eligibility Report. The SER is the official mode of communication between the processing center and the applicant. The SER notifies the applicant about his or her eligibility status and carries computerized messages about the application. The student must submit a signed SER to his or her financial aid administrator before a Basic Grant can be awarded.
Solicited Correction:	Refers to a change of application information on the transaction just subsequent to receiving a processing system edit.
Suspect Fields:	Refers to the application items which caused the applicant to be selected from validation. All of the PEC Group A sub-criteria were designed to elicit a correction to one or more fields.
Tax Filing Status:	Refers to the application item which asks applicants if (1) tax return figures are based on a completed return, (2) if those figures are estimated, or (3) if a tax return will not be filed for 1978.
TP:	Taxes paid. TP is an application item which refers to the amount of Federal Income Tax paid for 1978.
UE:	Unusual expenses. UE is the sum of two application items: (1) the amount of medical and dental expenses not covered by insurance plus (2) the amount of loss due to theft, or damage by fire, storm, or accident not covered by insurance for 1978.
UT:	Unreimbursed tuition. UT is an application item which refers to the amount of elementary, junior high, or high school tuition or fees paid for dependents in the household during 1978.
Unsolicited Correction:	Refers to a change of application information that is not in response to a processing system edit.

Validation: Applicants who were selected by the processing system to present certain documents to their financial aid administrators which confirm the accuracy of the information on their application form.

VEB: Veteran's educational benefits. VEB is an application item which refers to the monthly amount of educational benefits that the applicant expected to receive from the Veteran's Administration during the 1979-1980 academic year.